Faculty of Engineering Science
Department of Mechanical Engineering
Celestijnenlaan 300A
B-3001 Leuven Belgium

In cooperation with
imec
Kapeldreef 75
B-3001 Leuven Belgium

THERMO-MECHANICAL STUDY OF CHIP-PACKAGE INTERACTION EFFECTS FOR 3D STACKED IC TECHNOLOGIES

Members of the Examination Committee:
Prof. Pierre Verbaeten, chairman
Prof. Dirk Vandepitte, supervisor
Prof. Ingrid De Wolf, co-supervisor
Prof. Robert Pierre Mertens
Prof. Bert Verlinden
Prof. David Moens
Dr. Bart Vandevelde
Prof. Adrijan Barić (University of Zagreb, Croatia)

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor in Engineering Science

Uitgegeven in eigen beheer, ANDREJ IVANKOVIC, FACULTEIT INGENIEURWETENSCHAPPEN, DEPARTEMENT WERKTUIGKUNDE, AFDELING PMA

ISBN 978-94-6018-886-2
D/ 2014/7515/110

Acknowledgments

I would like to dedicate this PhD thesis to my mother and my grandfather. Dear Mom, throughout many years, you were always there for me and enabled me to be the person I am today. Thank you for the love, energy, constant support and finally the freedom you gave me to choose my own paths in life. I am grateful to be raised as I was. To my grandfather, who highly valued education and effort put in it, thank you for always believing in me and the respect I felt from you during all those steps.

The last 4 years were a tremendous journey. Like every true journey, it puts you through the hardest tests and gives you unmatchable experiences and rewards. There were sacrifices and turmoil and there was happiness and magic. There were places and people that I will never forget, many of those people with who I will always have a bond with. During this time, I grew as a person and engineer. Looking back, the bad times vanish and become only a valuable learning experience, while good times burstingly take over and put a silent smile on my face along with a feeling of deep content. After all, a ship in harbor is safe, but that is not what ships are built for. There are many people I would like to acknowledge, that one way or the other made an impact on me in this period. Please forgive me if a name ends up being ommitted, it is certainly not intentional.

Firstly a thank you to Bart Vandevelde, my first daily supervisor, for showing initial interest and giving me the chance to pursue this doctorate in the first place. Thank you for all the later support, always open doors and tensionless responsiveness. Thank you to my promotor Dirk Vandepitte, for accepting me as a candidate and the kind communication you exercized throughout all that time. Thank you to my co-promotor and department head Ingrid De Wolf for the technical and organizational support, the advices and guidance during all this time. A special mention goes to Mario Gonzalez and Vladimir Cherman. Mario, you exercized a very strict attention to results and reports always with the best intentions and entered numerous discussions with me. I have learned a lot from these moments and am grateful for your effort. Apart from the technical side, with your cheerful character and communicative nature you always created a great atmosphere and showed how important it is to have a positive attitude. Vladimir, without your practical knowledge, this thesis would not exist in the form it does today. Thank you for your patience, integrity, your immediate responsiveness to all those calls from the lab, for sharing your practical skills with me and for all technical and non-technical discussions in the office.

A thank you and appreciation goes to the supervisory and doctoral committee members for tracking my work, providing useful feedback and finally thoroughly evaluating it: Prof. Pierre Verbaeten, Prof. Dirk Vandepitte, Prof. Ingrid De Wolf, Prof. Bert Verlinden, Prof. Robert Pierre Mertens, Prof. Adrijan Barić, Prof. David Moens and Dr. Bart Vandevelde.

I was priviliged to be in a very special department at imec, where the atmosphere was always very friendly, where mutual respect was always felt regardless of position or experience and colleagues were not just technically competent but easily approachable, responsive and open minded. Thank you REMO group!

A person's foundations are built by his family and this thesis has a direct link to them as well. Dad, thank you for standing by my side, always asking about my prob-

lems, for giving advice, providing motivation when things were going slow and showing by your own example what it takes to get the job done properly. To my grandmas and grandpas, thank you for always welcoming me with your arms wide open, eagerly listening about my new experiences and being excited with me for all new steps made. Marin, thank you for teaching me the value of freedom, honor and trustworthiness through particular deeds.

I wouldn't be able to be a happy person without my friends. Behind each of the following names, there is a special story and for each and every one of these people I have a feeling of gratitude and respect. Vladimir, Matija, Raul, Neven, Dino, Sebo...14 years and going strong! Nabi, Anđelo, Alicja, Biljana, Luka, Federica, Andrea, Vice and Anna-Maria...thank you for the time spent together and all the memories. Let's continue to make them!

To all the people mentioned here and to many more that I have in mind but could not fit in this paragraph only due to limited space, thank you and all the best in your future endavours!

Abstract

Advances in performance of integrated circuits has throughout history been linked to minituarization of front-end-of-line (FEOL) devices, such as transistors. Three dimensional stacked integrated circuits (3D SIC) present a potential technology that can enhance system performance based on interconnect and packaging techniques. Contemporary vertical stacking of Si chips in 3D SIC technology is based on through-Si vias and microbumps. Compared to two dimensional integrated circuits, these vertical interconnects enable shorter signal paths that lead to lower signal delay and reduced power consumption. Extensive research is being done on the electrical properties of TSVs and microbumps. Furthermore, the thermo-mechanical impact of TSVs on the back-end-of-line (BEOL) and Si is being investigated. However, little is known about the thermo-mechanical impact of chip-package interaction (CPI) on Si in 3D SIC technology. Therefore, this thesis focuses on detecting mechanical stress mechanisms occuring after 3D IC stacking and packaging and the impact of those stresses on FEOL devices in Si.

Within this thesis, the linear piezoresistance model, originally developed for bulk Si, is applied to describe the response of FEOL devices to stress. In this case, the piezocoefficients represent the stress sensitivity of the whole device, rather than bulk material.

A stress sensor evaluation methodology is proposed and implemented within the thesis. Stress sensitivity of several FEOL device types and through several FEOL technology nodes was assessed. FEOL devices were calibrated to in-plane and out-of-plane stress. The stress sensitivities are analyzed from two perspectives:

- To gain information on the nature of stress sensitivity of FEOL devices and its trends through scaling technology nodes
- Their applicability for usage as FEOL CPI stress sensors

Obtained FEOL device piezocoefficients can greatly differ from Smith piezocoefficients for bulk Si. Direct calibration of FEOL devices to stress is necessary.

The stress sensitivity of MOSFETs, FinFETs and pseudo-Hall devices was assessed. All of these devices exhibit a sensitivity to mechanical stress. MOSFETs were implemented as stress sensors. An important advantage of MOSFETs is its ability to extract individual stress components. FinFETs exhibited lower sensitivity to stress than MOSFETs. Pseudo-Hall devices exhibit better temperature stability than MOSFETs, but extraction of stress components is not straightforward. Under certain conditions, MOSFETs can be used to extract both in-plane stresses and out-of-plane stress.

High MOSFET current shift of above 40% was observed after 3D Si die stacking and linked to the underfill-microbump stress mechanism. During cooling of a 3D IC stack after die bonding, the underfill as the material with the highest coefficient of thermal expansion (CTE) in the surrounding, shrinks and pulls the thin Si die over the underlying microbumps causing local warpage of the thin Si die. The underfill microbump stress mechanism is a CTE mismatch driven mechanism initiated by the underfill material.

The underfill-microbump interaction was experimentally monitored on 3 generations of 2-die 3D stacks. Global and local in-plane and out-of-plane stresses generated after 3D stacking were extracted with MOSFETs on both die levels. Stress in Si was monitored on the opposite side of the microbumps, directly below the microbump and between microbumps. After 3D die stacking, on the opposite side of the microbump, tensile in-plane stress components in Si are dominant. Compressive out-of-plane stress becomes dominant on the Si side below the microbump. Cu pads used for creating microbumps leave a non-negligible stress mark in Si as well.

N-type and p-type Si and in that sense n-type and p-type MOSFETs have a distinctively different response to mechanical stress. Circular stress patterns are observed in Si as a consequence of the underfill-microbump interaction, which reflect into circular current shift patterns for n-type Si devices and orbital current shift patterns with positive and negative current shift regions for p-type Si devices. Keep-out zones, prohibited areas in Si for processing MOSFETS and other sensitive FEOL devices on IC layouts due to the underfill-microbump stress impact, are proposed. Circular keep-out zones, primarily above the microbump position are proposed for n-type devices. Rectangular keep-out zones around the microbump position, in the surrounding but not directly above the microbump position, are proposed for p-type devices.

Underfill-microbump stress mitigation guidelines are proposed based on an experimentally validated finite element model. Increasing die thickness, choosing a low-stress underfill, decreasing the pitch between microbumps, decreasing the microbump height and grouping microbumps in larger arrays all contribute to stress reduction in Si. Based on implementation of some of the guidelines, such as increase of Si thickness and lower microbump pitch, stress decrease in Si is observed through assessed 3D stack generations, from a starting 827 MPa down to 125 MPa.

3D stacks were further packaged. Global and local in-plane and out-of-plane stress in Si generated after 3D stack packaging was extracted with MOSFETs on both die levels. It was revealed that the package substrate has an equally important impact in generating Si stress and package warpage as the mold compound. The final package warpage and Si stress is a result of the interaction of the substrate and mold compound.

Packaging exerts different global stress patterns on the two stacked Si dies. Between the microbumps, compressive in-plane stress is observed on the thinned die and tensile in-plane stress is observed on the thicker die. High out-of-plane stress is observed between the microbumps on the thicker Si die. Local sensors revealed that stress patterns in Si from the underfill-microbump reaction remain after packaging and are only modified in absolute value by the interaction of the substrate and mold compound.

Nederlandse Samenvatting

Doorheen de geschiedenis van geïntegreerde schakelingen is de vooruitgang van hun prestaties steeds gekoppeld geweest aan de miniaturisatie van componenten zoals transistors. Drie-dimensionele opeengestapelde geïntegreerde schakelingen (3D SIC) vertegenwoordigingen een technologie die de prestatie van het systeem kan verhogen op basis van technieken van verbindingen en verpakkingen. De moderne verticale opeenstapeling van Si chips in de 3D SIC technologie is gebaseerd op verbindingen doorheen silicium (TSVs) en micro-bumps. In vergelijking met twee-dimensionele geïntegreerde schakelingen, zijn deze verticale verbindingen korter wat leidt tot een kleinere signaalvertraging en een lager vermogensverbruik. Er wordt uitgebreid onderzoek gedaan naar de elektrische eigenschappen van TSVs en micro-bumps. Daarnaast wordt ook de thermo-mechanische impact van TSVs op Si onderzocht. Er is echter weinig gekend over de thermo-mechanische impact op Si van de interactie tussen de chip en de verpakking (CPI) in de 3D SIC technologie. Daarom richt deze thesis zich op het detecteren van de mechanismen van mechanische spanning na de 3D SIC opeenstapeling en de verpakking, en de impact van die spanningen op de componenten in het Si.

In deze thesis wordt het model van lineaire piëzo-resistentie, dat oorspronkelijk ontwikkeld werd voor bulk Si, toegepast om het spanningsgedrag van componenten te beschrijven. In dit geval vertegenwoordigen de piëzo-coëfficiënten de spanningsgevoeligheid van de gehele component. In de thesis wordt een evaluatie-methodologie voor spanningssensoren voorgesteld en geïmplementeerd. De spanningsgevoeligheid van verscheidene componenten doorheen verschillende technolgie-nodes werd bepaald. De componenten werden gekalibreerd voor spanningen in het vlak en uit het vlak. De spanningsgevoeligheid werd vanuit twee perspectieven geanalyseerd:

- Om informatie in te winnen over de aard van de spanningsgevoeligheid van componenten, en de verandering ervan door schalen van de technologie-node
- Hun toepasbaarheid voor het gebruik als CPI spanningssensoren

De verkregen piëzo-coëfficiënten voor de componenten kunnen sterk verschillen van de piëzo-coëfficiënten van Smith voor bulk Si. Directe kalibratie van componenten aan spanning is noodzakelijk.

MOSFETs, FinFETs en pseudo-Hall componenten zijn gevoelig voor mechanische spanning. MOSFETs worden gebruikt als spanningssensoren. Een belangrijk voordeel van MOSFETs is het feit dat ze toelaten om individuele spanningscomponenten te verkrijgen. FinFETs vertoonden een lagere gevoeligheid voor spanning dan MOSFETs. Pseudo-Hall componenten vertonen een betere thermische stabiliteit dan MOSFETs, maar het is niet vanzelfsprekend om de spanningscomponenten te verkrijgen. Onder bepaalde voorwaarden kunnen MOSFETs gebruikt worden om zowel de spanning in het vlak als uit het vlak te verkrijgen. Na 3D-stapelen werd een hoge verschuiving van meer dan 40% van de stroom van de MOSFET waargenomen, die gelinkt werd aan het underfill-micro-bump spanningsmechanisme. Tijdens het afkoelen van een 3D-stapeling na het verbinden van de chips, krimpt de underfill als het materiaal met de hoogste CTE in de omgeving, en trekt het dunne stukje Si over de onderliggende micro-bumps, wat lokale buiging van het dunne stukje Si veroorzaakt.

Het underfill-micro-bump spanningsmechanisme is gedreven door een discrepantie in de thermische uitzettingscoëfficiënt, geïnitieerd door het underfill-materiaal.

De interactie tussen de underfill en de micro-bumps werd experimenteel gevolgd voor 3 generaties van 3D-stapelingen met 2 niveaus. De globale en lokale spanning in het vlak en uit het vlak ontstaan na 3D-stapeling werden verkregen met MOSFETs op beide niveaus. De spanning in het Si werd gevolgd aan de zijde van het stukje Si tegenover de micro-bumps, direct onder de micro-bumps en tussen micro-bumps. Na 3D-stapeling zijn, aan de zijde tegenover de micro-bump, trekspanningen in het vlak dominant in het Si. Drukspanningen worden dominant aan de zijde van het Si onder de micro-bump. Ook de Cu pads gebruikt om de micro-bumps te creëren, laten een niet-verwaarloosbaar spoor in het Si.

N-type en p-type Si en in die zin n-type en p-type MOSFETs hebben een opvallend verschillende respons op mechanische spanning. Cirkelvormige spanningspatronen worden waargenomen in Si als gevolg van de interactie tussen de underfill en de micro-bump, die zich weerspiegelen in cirkelvormige patronen in de stroomverschuiving voor n-type Si-componenten en in orbitale patronen in de stroomverschuiving met gebieden van positieve en negatieve verschuiving voor p-type Si-componenten. Veiligheidszones, verboden gebieden in Si voor het plaatsen van MOSFETs en andere gevoelige componenten omwille van het effect van de underfill-micro-bump spanning, worden voorgesteld. Voor n-type componenten worden cirkelvormige veiligheidszones voorgesteld, voornamelijk boven de positie van de micro-bump. Voor p-type componenten worden rechthoekige veiligheidszones rond de positie van de micro-bump voorgesteld, in de omgeving van maar niet direct boven de positie van de micro-bump.

Op basis van een experimenteel gevalideerd eindige-elementenmodel worden richtlijnen voorgesteld voor het verlagen van de mechanische spanning. Factoren die bijdragen tot de verlaging van spanning in het Si zijn een toename van de dikte van de stukjes Si, de keuze voor een underfill met lage spanning, het verlagen van de afstand tussen micro-bumps, verlagen van de hoogte van de micro-bump en het groeperen van micro-bumps in grotere reeksen. Door de implementatie van sommige van deze richtlijnen, zoals een toename van de dikte van het Si en een kleinere afstand tussen de micro-bumps, wordt een spanningsafname van 827 MPa tot 125 MPa vastgesteld doorheen de geëvalueerde generaties van 3D-stapelingen.

De 3D-opeenstapelingen werden verder verpakt. De globale en lokale spanning in Si, in het vlak en uit het vlak, gegenereerd na verpakking van de 3D-opeenstapeling werd verkregen met MOSFETs op beide niveaus van stukjes Si. Het substraat van de verpakking heeft een even belangrijke invloed op het genereren van spanning in het Si en de buiging van de verpakking, als de mold compound. De uiteindelijke buiging van de verpakking en de spanning in het Si zijn het resultaat van de interactie van het substraat en de mold compound. Het verpakken oefent verschillende globale spanningspatronen uit op de twee gestapelde stukjes Si. Tussen de micro-bumps wordt een drukspanning in het vlak waargenomen op het verdunde stukje Si en een trekspanning in het vlak op het dikkere stukje Si. Er wordt een hoge spanning uit het vlak waargenomen tussen de micro-bumps op het dikkere stukje Si. Lokale sensoren lieten zien dat de spanningspatronen van de interactie tussen de underfill en de micro-bumps behouden blijven na verpakking, en enkel in absolute waarde gewijzigd worden door de interactie tussen het substraat en de mold compound.

List of acronyms

3D(S)IC	3-dimensional (stacked) integrated circuit
4PB	4-point bending
BCB	benzocyclobutene
BEOL	back-end-of-line
BGA	ball-grid array
CBCM	charge-based capacitance measurement
CIEF	charge-injection-induced error-free
CPI	chip-package interaction
CTE	coefficient of thermal expansion
DCB	double cantilever beam
DCM	dynamic contact module
DCU	Dublin City University
DRAM	dynamic random-access memory
FEM	finite element method
FEOL	front-end-of-line
FinFET	fin field-effect transistor
FWHM	full width half maximum
MCM	multi-chip module
MEMS	micro-electro-mechanical system
MOSFET	metal-oxide-semiconductor field-effect transistor
NUF	no-flow underfill
PCB	printed circuit board
PSI	phase-shifting interferometry
RC	rocking curve
SiP	system in package
SMU	source-measurement unit
SoC	System-on-Chip

TSV	through-Si via
ULK	ultra low-k
UTCS	ultra-thin-chip-scale
UTS	ultimate tensile strength
VASE	Variable Angle Spectroscopic Ellipsometry
VSI	vertical scanning interferometry
WLP	wafer level package
XRD	X-ray diffraction
3DSM	3D surface modeling

List of Figures

List of Tables

Contents

Chapter 1

Introduction

The introductory chapter lays the setting for this PhD topic. It mentions its field of study, lays out the thesis objective and briefly lists the content of each chapter.

1.1 Thesis objective

This PhD thesis finds its place within a particular field of microelectronics - three dimensional stacked integrated circuits (3D SIC). This work was performed from 2010-2014 at imec, Belgium as a part of the 3D integration program.

The objective of this PhD thesis is to detect, describe and provide solutions for thermo-mechanical effects of 3D IC stacking and packaging on the Si die front-end-of-line (FEOL) in order to improve 3D IC stack/package performance. This objective is channeled into the following main goals:

- Develop a sensor-based chip-package interaction (CPI) methodology for the FEOL
- Provide understanding and mitigation guidelines on the generated mechanical stress to the FEOL in the vicinity of the microbumps after 3D stacking - referred to as the underfill-microbump stress mechanism

Additional goals developed throughout the thesis include:

- Enable benchmarking of FEOL device stress sensitivities from scaling technology nodes by applying stress calibration methodology used for stress sensors
- Monitor and interprete stress evolution after packaging of 3D IC stacks

1.2 Thesis outline

The following section gives an overview of the content of each chapter in this thesis. Each chapter in this work is preceeded by a short abstract and followed by a summary.

Chapter 2 starts with a historical overview of 3D stacking of electronics leading to its role in todays microelectronics industry. Thereafter, the contemporary technical surrounding in which this thesis commenced is mentioned. Further on, a literature overview related to 3D CPI is given including a section on FEOL stress sensors. The chapter concludes by outlining the original contributions of this PhD thesis.

Chapter 3 initially describes the basics of the piezoresistance effect in Si and its application to this work. This is followed by a presentation of the methodology for stress sensor evaluation.

Chapter 4 lists all of the utilized analysis methods and techniques. These include the finite element method (FEM), 4-point bending, nano-indentation, optical and mechanical profilometry and X-ray diffraction.

Chapter 5 presents a study on impact of stress to FEOL devices, from two points of view. The first one being evaluation of stress sensors for usage in 3D stacking and packaging applications and the second one benchmarking FEOL device sensitivities with scaling from the 130nm technology node to the 32 nm technology node. The devices characterized include MOSFETs, FinFETs and pseudo-Hall transistors.

Chapter 6 discusses a back-end-of-line (BEOL) out-of-plane stress sensor, as a possible alternative to FEOL out-of-plane stress sensors.

Chapter 7 incorporates an in-depth analysis of the underfill-microbump stress mechanism observed after 3D IC stacking. It covers fundamental understanding of

the mechanism's origin and monitoring its effect in the FEOL through 3 generations of 3D stacks. Furthermore, the impact of the underfill-microbump stress mechanism to IC layouts is discussed and guidelines for mitigation of its impact on the FEOL are proposed.

Chapter 8 continues with an analysis of mechanical stress in the FEOL after packaging 3D stacks. Stress was monitored in realized 2D and 3D packages, with particular focus on local stress, related to the evolution of the underfill-microbump interaction, and global stress across the die.

Chapter 9 summarizes the main findings in this thesis and proposes future work.

Chapter 2

State of the art and beyond

Chapter 2 gives insight into the evolution of 3D integrated circuits through history and summarizes studies related directly and indirectly to chip-package interaction (CPI), until the commencement of this thesis. The introduction to 3D integrated circuits begins with revealing its origins and first ideas of stacking electrical circuits in the 1950s in section 2.1.1. The evolution of stacked electrical circuits and subsequently stacked integrated circuits in the following decades is covered in section 2.1.2. Section 2.2 gives an overview of available literature related to CPI in 3D IC technologies. Section 2.2.1 focuses on 3D stacking and packaging while section 2.2.2 gives an overview on mechanical stress sensors used to extract Si stress in microelectronics. Section 2.2.1 is further divided into 3 subsections. Section 2.2.1.1 discusses the thermo-mechanical impact of the through-Si via (TSV) in Si dies, often not included in chip-package interaction topics depending on considering the TSV an integral part of the die or a stacking/packaging assembly step. Following the thermo-mechanical impact of TSVs in Si, several 3D IC stacking methods studied until this point are presented in section 2.2.1.2. Section 2.2.1.3 summarizes work focusing on non-TSV related CPI topics in 3D ICs. Finally, after the historical and literature overview of 3D integrated circuits, the contributions of this thesis are stated.

2.1 Brief history of 3D integrated circuits

2.1.1 Idea of verticality

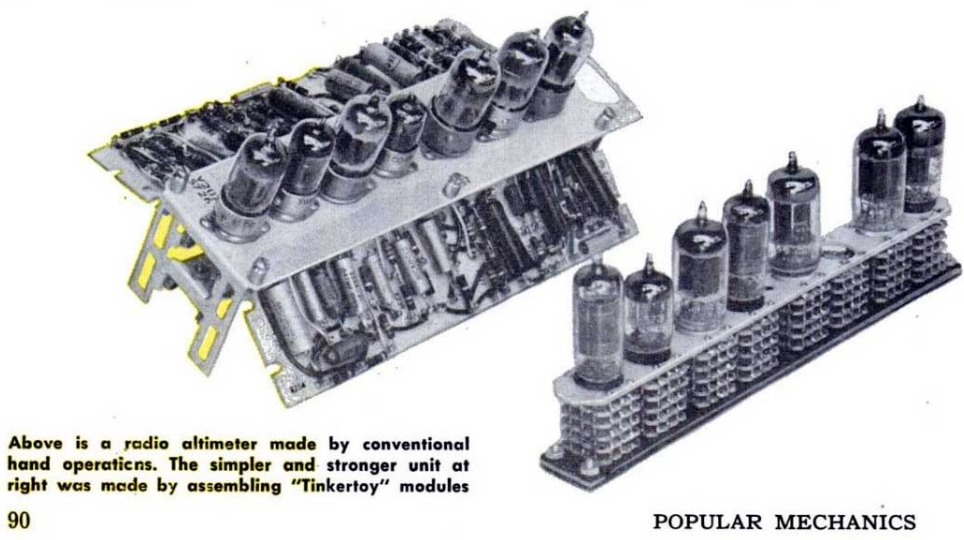
All engineering disciplines have several common goals they aspire to embody in their systems. Some of the more prominent ones are certainly efficiency of the system, its speed and its size. Achieving these qualities lies within the scope of the creativity of the engineers and currently available technology. The first idea of vertically stacking electronics came already in 1951 from the US Navy. What was then called "Project Tinkertoy", envisioned one-component 0.625 inch wafers stacked on each other for the sake of less consumed space and was thought would create a more tough and reliable unit. A black and white photograph of the "Tinkertoy" 3D stack is shown in figure 2.1 [1]. The conventional planar approach to circuit design, on the left hand of the photo, was transformed to the vertical stack, on the right side the photo. It took another couple of years for higher interest and wider implementation of this idea.

In 1957, the Radio Corporation of America developed a 3D stacked electrical circuit which was in later years extensively implemented in radio based devices and started gaining commercial popularity. These stacks were made with 0.36-inch ceramic wafers, twice as smaller than the "Tinkertoy" and were connected by riser wires soldered to solder ridges at the edge of the wafer. The entire module could be encapsulated and when assembled stood only 0.4 to 0.8 inch tall. Due to their minimalistic dimensions they were named Micromodules. Figure 2.2 [1] depicts a multilevel Micromodule stack compared to a match, with visible interconnections on the sides of the structure leading to the base. Interesting enough, although made 50 years ago, these stacks visually resemble today's idea of 3D stacking technology. Figure 2.3 [2] presents assembled Micromodules on a PCB and a direct showcase of floorplan optimization coming with it.

After millions of testing hours, the Micromodules proved to have soaring reliability with RCA eventually noting they are 6 times more reliable than conventional transistor based planar circuits of the time and 60 times more reliable than vacuume tube circuits, still used in some industry segments then. The interest for Micromodules rose to 60 companies established within a couple of years specializing in their production for specific electronics industry branches.

However, regardless of their advantages and hype they created in a short period of time, micromodule development soon unexpectedly came to a halt. The event that eventually lead to the downfall of Micromodules is the production of the first integrated circuit by Jack Kilby of Texas Instruments in 1959. With devices capable of being integrated in silicon, the space consumption increased even more with immense potential of technology scaling and further developement in speed and power. By 1964 the industry focus changed completely back to planar production which now made a turn to integrated semiconductor circuits and started the era of system-on-chip (SoC) arhitectures. The exact same year, IBM produced one of the first commercial logic integrated circuit for their business computer line, the Solid Logic Technology (SLT), presented in figure 2.4 [3].

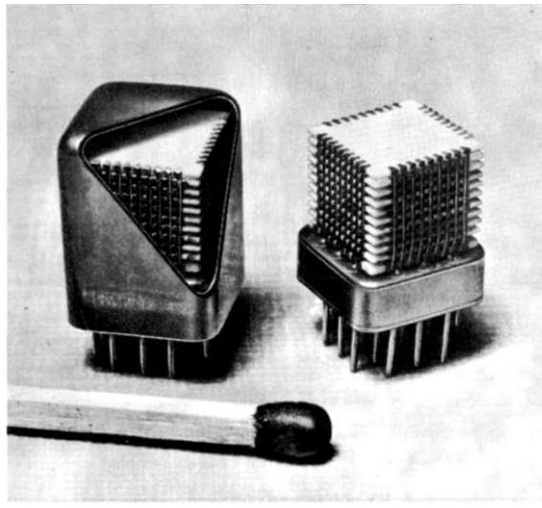
It is interesting to mention that at the peak of Micromodule production in 1962 Earle F. Cook, chief signal officer of US army, then the main consumer of Micromod-



Above is a radio altimeter made by conventional hand operations. The simpler and stronger unit at right was made by assembling "Tinkertoy" modules

90 POPULAR MECHANICS

Figure 2.1: Project "Tinkertoy" 3D stack [1]



Micromodules with electron beam welded interconnections

(Courtesy: Hawker Siddeley Dynamics)

Figure 2.2: Micromodules 3D stack [1]

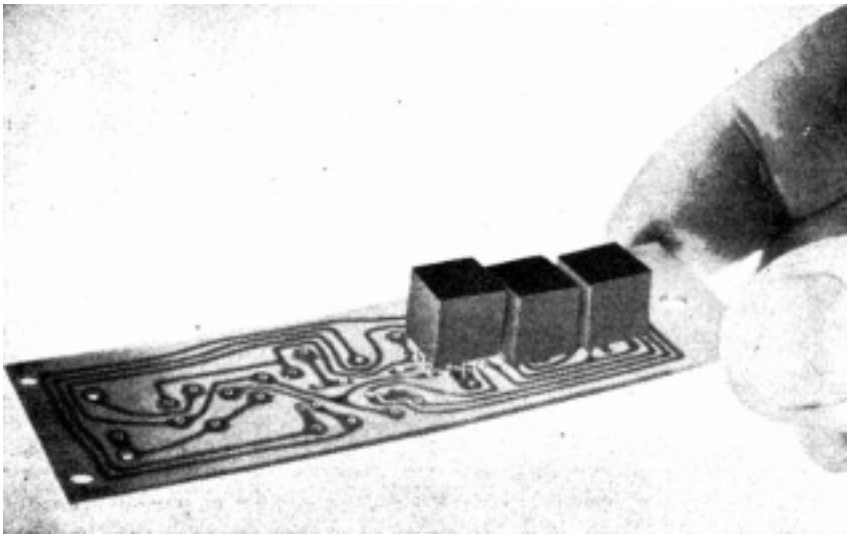


Figure 2.3: Encapsulated Micromodules embedded on a PCB [2]

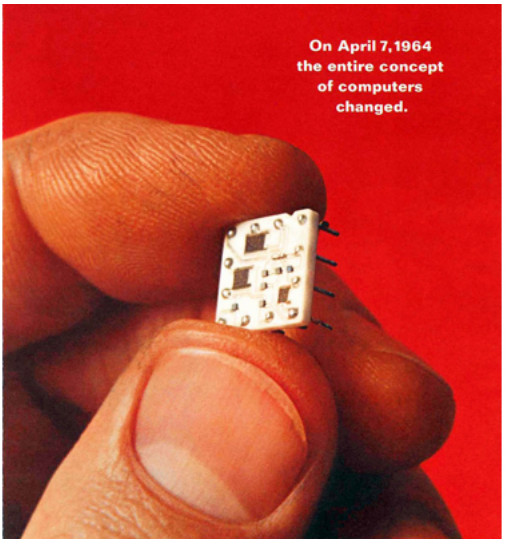


Figure 2.4: One of the first commercially available logic ICs, IBM's Solid Logic Technology [3]

ules, made a statement where he foresaw integrated circuits and 3D stacking going hand-in-hand and joining forces for future technologies. Cook noted that the micromodules would accommodate advanced integrated circuitry, "logically, through evolution, rather than by revolution. Such developments normally mature bit-by-bit over a protracted period of time, rather than suddenly appearing as an entire operational system. The industry anticipates a long period of applications of the new devices in hybrid combinations with conventional components" [2]. The micromodule program was however completely shadowed by ICs, with 3D stacking type developments left only for specific applications, out of commercial focus until the first years of the 21st century, when the idea was fully renewed on a larger scale.

2.1.2 Evolution and diversity of 3D IC technology

The idea of verticality of circuits never vanished completely due to its attractive potentials for size and performance enhancement and attempts have been made in the meantime to apply the idea of verticality to integrated circuits. Studies on 3D integration of integrated circuits are consistently appearing again from the 1980s [4] and through the 1990s [5] with different ideas for stacking methods. Several stacking technologies have been proposed ranging from highly integrated monolithic stacks [6-8] to semi-integrated wafer level packaging (WLP) methods [9-11] and 3D packaging methods with low integration that mechanically place chips on top of each other but interconnect them off-chip, such as stacked Multi-Chip Modules (MCM) [12] and System in Package (SiP) types [13-15]. Monolithic stacks to this day seemed to mostly remain on idea level due to the complex technology required to enable such processing and interconnects. Stacked MCMs and in particular SiP became the most popular and in fact widely used 3D packaging approach due to its relative straightforwardness and simplicity, although still not frequently used in the commercial sector. Figure 2.5 [16] shows an example of a stacked MCM, where stacking is performed on module level. In this case, several packages are interconnected within one PCB and the PCBs subsequently stacked on each other. Sophisticated stacking technology is not necessary here and level of integration compared to other stacking methods is minimal. Figure 2.6 presents the other forementioned stacking technologies: a) SiP, b) WLP and c) monolithic stacks [17].

SiP products consist of several packages acting as one. They can contain one or more dies on each package strata where the dies are connected to the surrounding by means of wirebonds or ball-grid arrays (BGAs). In the SiP illustration in figure 2.6 a) a combination of wirebonds and solder balls is present, where some dies are connected to their substrates with wirebonds, others with solder balls and the stratas are interconnected solely with solder balls. WLP in essence relates more to the particular method of stacking but it also reflects in the used technology to do so. In this case, stacking assembly processes are initiated already on wafer level and dicing is done only after the package has been completed. Figure 2.6 b) depicts imec's ultra-thin-chip-scale (UTCS) WLP, where large vias are used to interconnect vertical layers. Vias penetrating through Si can be present here as well but in much larger dimensions and usually resembling more to plate shaped interconnects rather than the later developed fully material filled through-Si vias (TSVs). Monolithic 3D integration approaches stacking on transistor

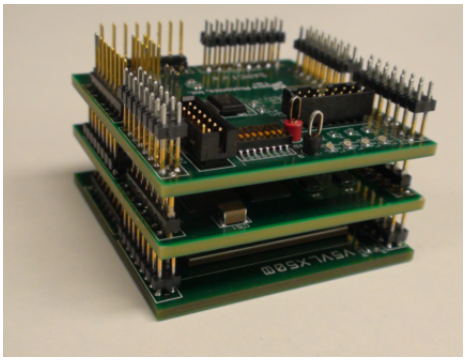


Figure 2.5: Stacked Multi-Chip Module [16]

level where only 100 nm or less of Si divides two stratas. The mismatch between the required high level of integration and limited processing technology prevent this 3D stacking approach to shift from prototypes to production. However, it still holds the high potential label, as perhaps one of the future steps of 3D IC technology.

With the turn of the century and SiP and WLP being the most popular 3D IC technologies that made it to the market due to their straightforward and relatively complex-free assembly process, the gap between harsh FEOL device scaling and interconnect and packaging development was still widening. The development of FEOL technology and the rapidly increasing amount of transistors per area has lead to communication on a single die progressing at an extremely fast pace. Figure 2.7 presents Intel's daily data exchange forecast from 2012 encompassing data starting from 2005. Daily data traffic can be already measured on exabyte scales and shows no signs of slowing down. The number, design complexity and subsequent length of interconnects being processed on one die was booming and it was clear that circuit speed and power consumption will soon be bottlenecked by interconnects rather than FEOL innovations.

The stage has been set for a new interconnect technology and the focus has turned to 3D stacked integrated circuits (3D SIC) with TSVs in focus. Research on TSVs began first in the 1990s and micrometers wide TSVs penetrating through Si showed indications for the potential to bring together both enhanced electrical performance and mechanical assembly feasibility. The introduced TSV based stacks provides shorter interconnects that bring faster speed and lower power consumption while the vertical architecture allows parallel data processing taking some weight off of heavy data traffic. Figure 2.8 illustrates a 3D SIC stack cross-section with several dies vertically connected with TSVs. As TSVs ensure only connection of two Si sides, other interconnects are also needed to link to other dies. More on 3D SIC, its features and problems will be discussed in section 2.2.

3D SIC can be considered a compromise regarding integration feasibility. A significant integration step forward from SiP and WLP to enable IC performance enhancement and a step back from monolithic 3D ICs to put processing feasibility and cost in a more realistic perspective.

In summary, 3D SIC technology has the potential to provide the following advance-

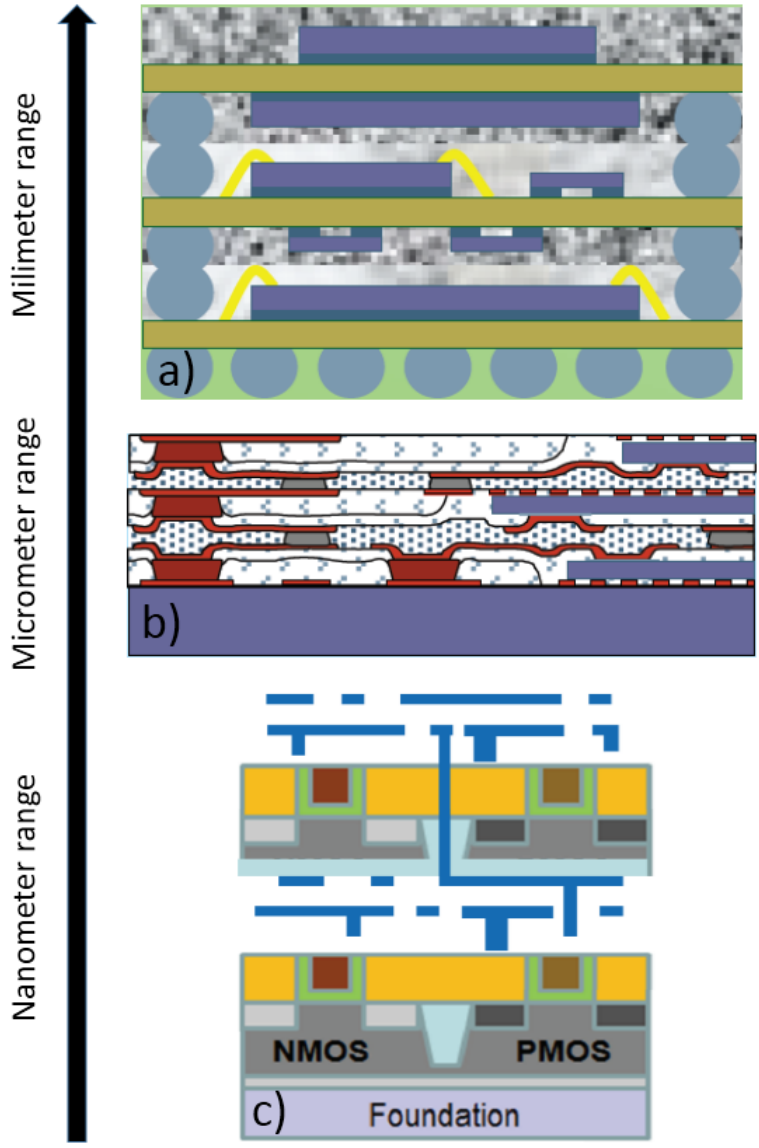


Figure 2.6: Contemporary stacking technologies: a) system in package (SiP), b) wafer level package (WLP) and c) monolithic stacks. An SiP ranges several mm in height, a WLP several hundreds of μm in height and monolithic stacks several hundreds of nm in height.

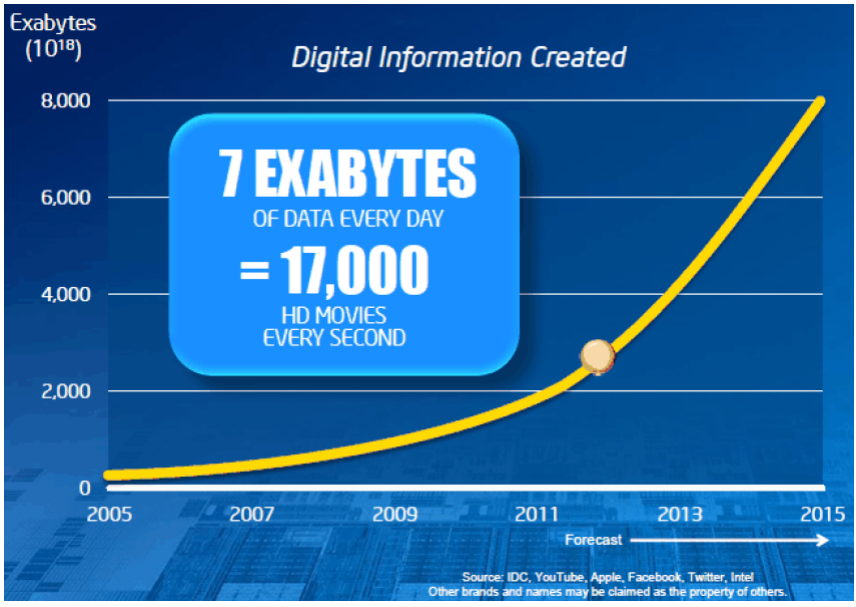


Figure 2.7: Intel's daily data exchange forecast from 2012

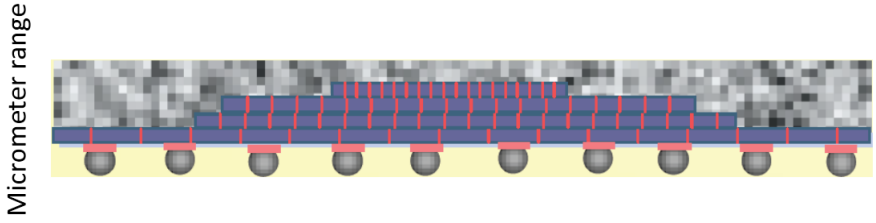


Figure 2.8: An illustration of a packaged 3D IC stack. The 3D IC stack with several Si dies can be below 100 μm or several hundreds of μm while the whole package is usually below 1 mm.

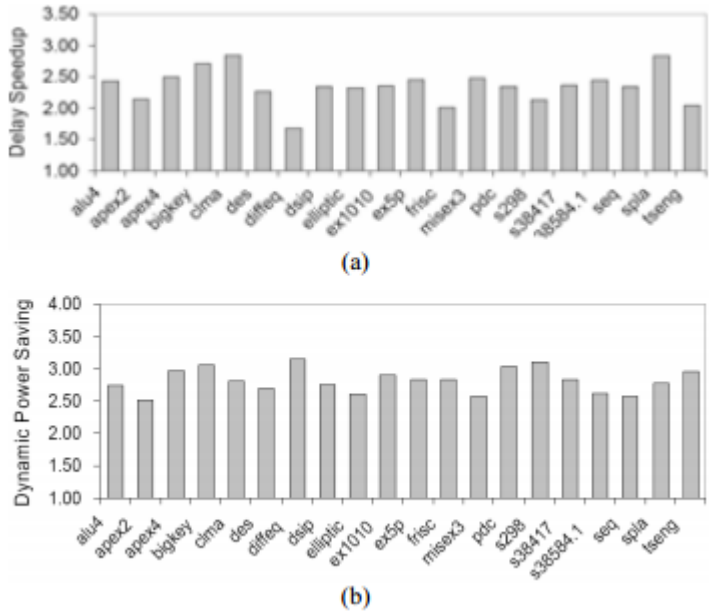


Figure 2.9: Study results presenting a) delay speed up and b) power savings on particular encrypted test circuits when embedded in 3D IC technology compared to 2D processing [18]

ments to microelectronics:

- Shorter interconnects leading to lower power consumption and lower signal delay and increased circuit speed (figure 2.9 [18])
- Smaller PCB footprint (figure 2.10)
- Heterogenous integration with a complete analog and digital system through various technology nodes, all within one package (figure 2.11)

This PhD thesis finds its place in the evolution of 3D ICs by focusing on chip-package interaction (CPI) thermo-mechanically induced phenomena in 3D SICs and their problematic impact on primarily the FEOL. From a wider perspective, this PhD participates in one additional attempt to use the benefits of 3D stacking and develop it for widespread commercial purposes. Only this time, 3D stacking does not seem just like an option, but a technical necessity. It gaines even more importance considering that few companies remain with funding capability for standalone investment in FEOL scaling based IC performance enhancement. Interestingly enough, 3D SIC technology research seems to have excelled coninciding with the 3D SIC stack prototype of the Pentium 4 processor produced by Intel in 2004 [19] acting perhaps as a kind of trigger directing focus on emerging technologies.

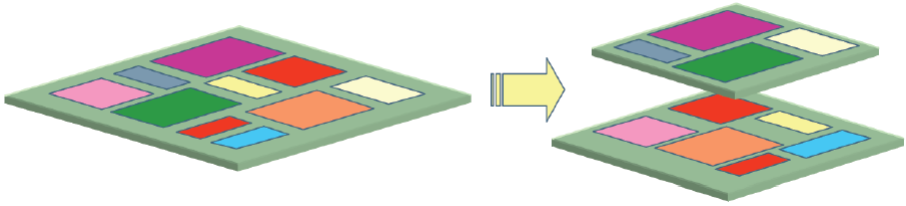


Figure 2.10: 2D to 3D transition - floorplan savings

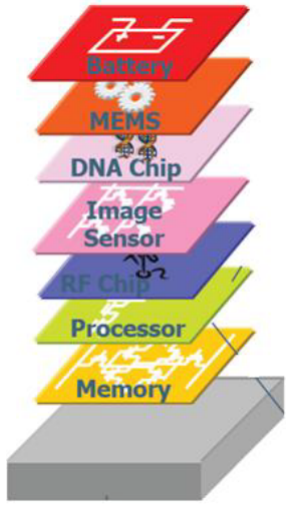


Figure 2.11: The potential of heterogenous integration within 3D IC stacks

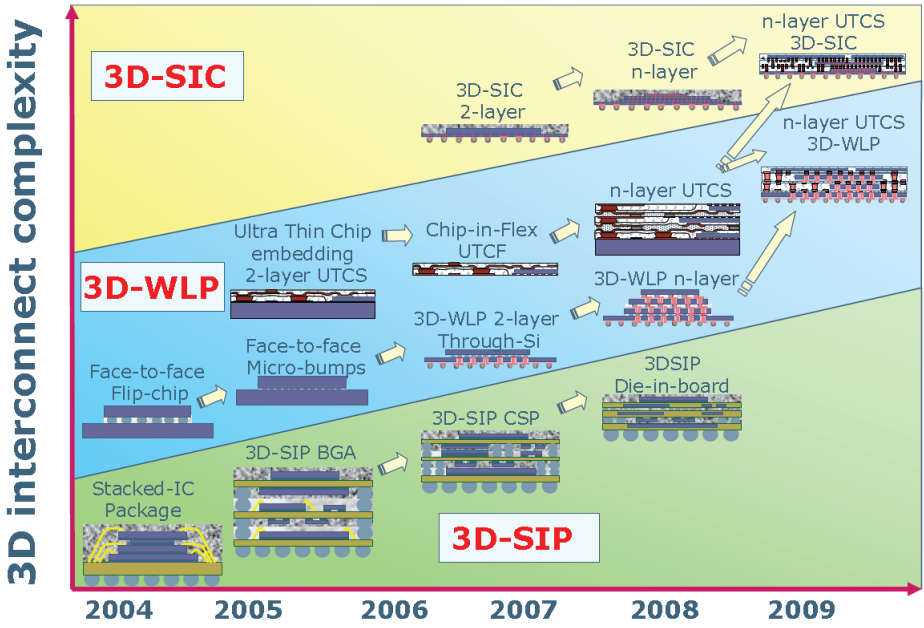


Figure 2.12: imec's 3D IC stacking/packaging roadmap from 2007

2.1.3 Contemporary setting of the thesis

While BEOL interconnects have been following FEOL scaling from the start of microelectronics, microelectronic packaging advancements exhibited a slower development slope, more often relying on older mature technologies. The 1990s brought first research on new packaging technologies. During early development of full TSVs, UTCS WLP packaging technology emerged, presented in figure 2.6 b) [10], which was more technology friendly at the time. Dies could be stacked in several layers and were interconnected with large vias. The dies were however not connected directly but with interconnect circuitry placed around the die in the surrounding wafer layers. The first turn to 3D SICs and beginning of publicly established 3D SIC programs is observed from 2004 with one of the first thermo-mechanical publications including TSVs seen in 2005 [20]. imec's 3D integration roadmap from 2004 to the beginning of this PhD in 2010 is presented in figure 2.12. It covered three types of 3D assembly: 3D SIP, 3D WLP and 3D SIC. Table 2.1 summarizes the main characteristics of the 3 technologies and quantifies their processing possibilities. Proposals were made for 3D SIP and 3D WLP with focus on UTCS technology but majority of resources were turned to 3D SIC, the representative of the shift form, with highest level of integration and most performance increase potential. 3D SIC technology presented the first stacking technology which starts already at Si foundry level.

Until the start of this PhD work in 2010, within the realm of 3D SIC, research focused on a large scope of topics that would enable first stacked prototypes. This includes TSV processing [21], initiation of studies related to TSV impact on the FEOL

	3D-SIP	3D-WLP		3D-SIC
Technology	Package interposer	WLP, Post-passivation		Si-foundry, Post FEOL
3D interconnect	Package I/O	UTCS Embedded die	Si-through vias	Si-through "Cu nail" vias
Intercon. Density	'package-to-package'	'around' die	'through' die	'through' die
Peripheral	2 - 3 /mm	10 - 50 /mm	10 - 25 /mm	25 -100 /mm
Area-array	4 - 11/mm^2	100 -2.5k/mm^2	16 - 100/mm^2	400-10k/mm^2
3D Si Via pitch	-	-	40 – 100 µm	< 10 µm
3D interconnect pitch	300 – 500 µm	20 – 100 µm	-	-
3D Si Via diameter	-	-	25 - 100 µm	1 - 5 µm
Die thickness	> 50 µm	10 - 20 µm	50 - 100 µm	10 - 20 µm

Table 2.1: Summary of main technical characteristics from the 3 contemporary 3D stacking technologies: SIP, WLP and 3D SIC

[22] and BEOL [23], technology of stacking and stack type options [24-27]. Some of the more prominent approaches followed until 2010 within 3D SIC follow.

Related to TSV processing::

- TSV first
- TSV middle

Related to stacking methods:

- Cu-Cu bonding
- Cu insertion bonding
- Microbumps

Related to stacking types:

- Die to die
- Die to wafer
- Wafer to wafer

In 2010, the 3D integration program within imec started assembling prototypes which involve a TSV middle approach and die to die stacks with Cu-Sn based microbumps in between. The first 3D SIC prototype within the program was produced in 2010 where a thick DRAM was stacked on a 25 um thin logic die. A basic overview and potential interest areas within the stack that were published at the time are shown in figure 2.13. This PhD commences with the introduction of this first 3D SIC prototype.

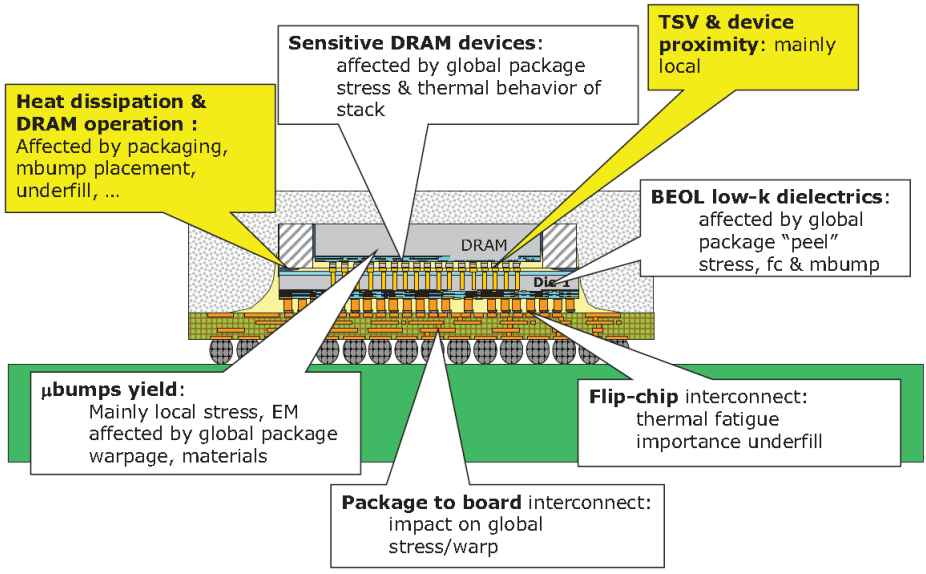


Figure 2.13: 3D SIC prototype in 2010 - DRAM on logic, coinciding with the start of this thesis

2.2 Literature overview on CPI in 3D IC technologies

This chapter defines chip-package interaction in 3D IC technologies and proceeds to a literature overview of CPI subtopics in the following way:

- Literature related to the topic of this PhD work, thermo-mechanically caused CPI effects in 3D ICs and stress sensors, will be given only until the starting date of this PhD work, in July 2010 in order to provide a clear impact of this PhD later on
- Literature for a subtopic closely linked to this PhD, but not its integral part, TSV impact on the FEOL, will be mentioned also beyond 2010 for an up-to-date overview on the topic
- CPI effects in 3D ICs which do not have a thermo-mechanical origin are left out in this overview

In general, chip-package interaction is a broad term and can cover any interaction between the chip and package which causes changes in chip/package performance or chip/package failures. In 2010, 5 reliability domains within 3D ICs were defined, presented in figure 2.14:

- TSV
- TSV and its environment

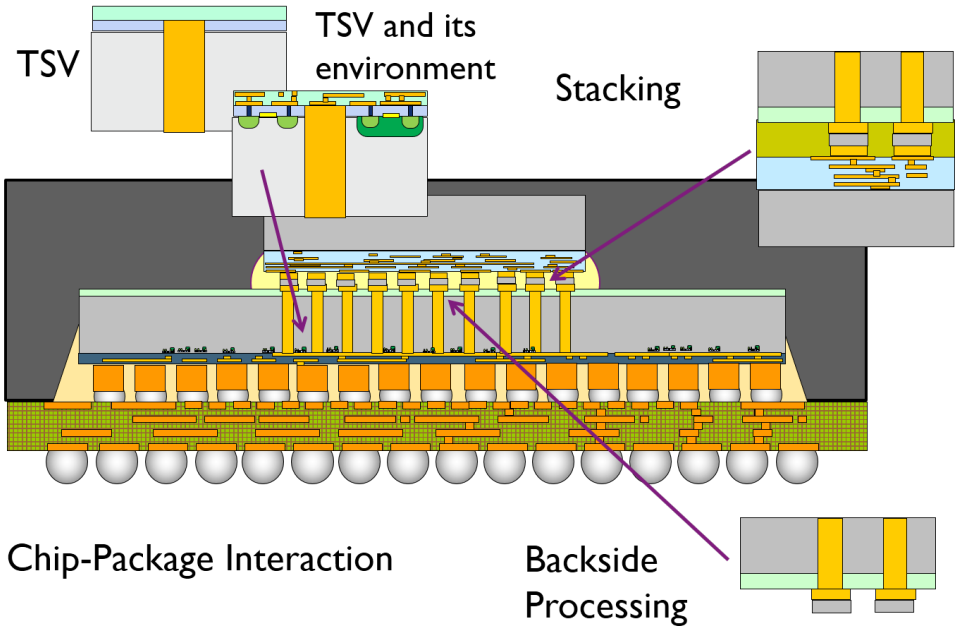


Figure 2.14: 5 reliability domains in 3D SIC technology defined in 2010 representing 3D SIC reliability topics needing immediate attention [I. De Wolf, K. Croes (imec)]

- Stacking
- Backside Processing
- Chip-Package interaction

TSV relates to the reliability issues of the TSV itself such as Cu instability, voids and barrier/liner integrity. The impact of the TSV on its environment, FEOL and BEOL, is encompassed in the second topic. Stacking and backside processing involve reliability issues related to wafer or die thinning and post-TSV assembly methods of Si stacking. Chip-Package interaction within 3D IC in a broad sense includes all interactions between the 3D package and interconnection constituents and the Si die. In most cases, this refers to thermal, thermo-mechanical or mechanical interactions within the package. This literature overview will focus on thermo-mechanical CPI effects, as this is the core topic of this PhD work. Stress sensor overview will be presented in a separate subchapter as stress sensor literature in 3D ICs is limited and due to the fact that stress sensors in 3D ICs are often built on previous 2D experiences.

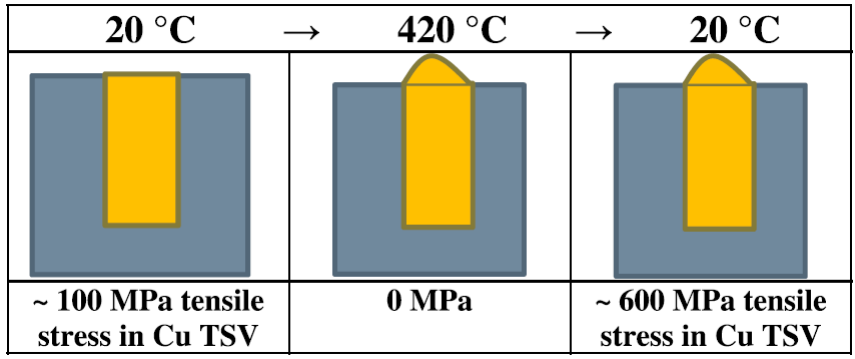


Figure 2.15: Plastic deformation of Cu after thermal treatment - Cu pumping. During BEOL processing Cu TSVs are exposed to high temperatures, the higher coefficient of thermal expansion (CTE) of Cu forces it to expand more than Si. This causes compressive stress in the confined Cu inside the TSV. This stress can partly be released near the top of the TSV, by out-of-plane expansion of the Cu, named Cu pumping, which can severely damage the BEOL layers. [23]

2.2.1 3D stacking and packaging

2.2.1.1 TSV impact

The impact of TSVs on their environment is at the borders of the CPI field. The TSV is a part of the Si die and impacts its surrounding locally, namely, the BEOL and FEOL. However, in essence the TSV is fundamental to 3D IC stacking and acts as a package level interconnect between dies or dies and the package. Some of the most commonly investigated thermo-mechanical impacts to the BEOL [23,28-31] and FEOL [22,32-35] will be mentioned here.

The TSV is a Cu based nail running through the Si material. Cu deforms plastically and has a CTE approximately 7 times higher than Si. During a series of thermal processes the TSV is submitted to during fabrication, the Cu and Si experience high material expansion and shrinking mismatches which can lead to effects such as Cu plastic deformation and material delamination.

Figure 2.15 taken from [23] illustrates the effect of Cu pumping where after thermal treatment Cu deforms plastically leaving a Cu bump above the TSV. This effect, in literature also referred to as Cu extrusion, Cu protrusion or Cu pop-out, causes deformation in the BEOL in the TSV surrounding. Figure 2.16 [23] provides a Cu pumping SEM image of the effect where the BEOL is seen bent following the TSV curvature. During BEOL processing Cu TSVs are exposed to high temperatures, the higher coefficient of thermal expansion (CTE) of Cu forces it to expand more than Si. This causes compressive stress in the confined Cu inside the TSV. This stress can partly be released near the top of the TSV, by out-of-plane expansion of the Cu, named Cu pumping, which can severely damage the BEOL layers.

Another effect under investigation is delamination of TSV from its surrounding.

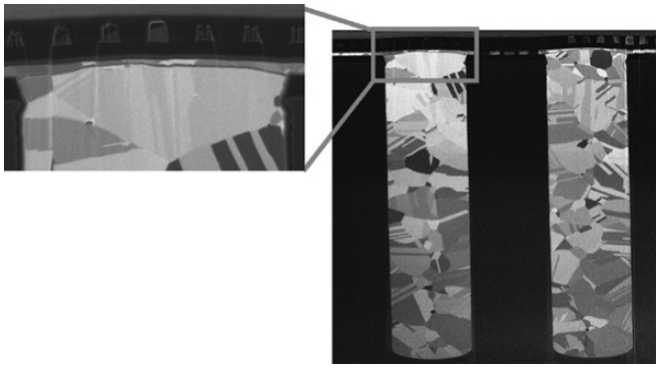


Figure 2.16: SEM image of the Cu pumping effect seen on a die cross section [23]

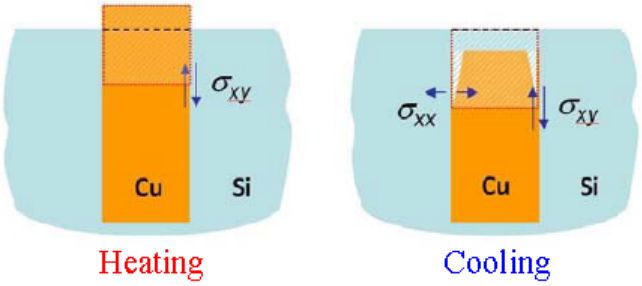


Figure 2.17: Stress components arising during TSV heating and cooling capable of initiating cracks. During heating shear stress occuring near the wafer surface will result in delamination. During cooling, the delamination is driven by both the shear stress and the radial tensile stress. [28]

During heating and cooling out-of-plane shear stress in combination with in-plane normal tensile stress in the TSV, referenced to the Si surface, can cause a mode II or mixed mode fracture initiating delamination of the TSV. Figure 2.17 taken from [28] illustrates the driving stress components during TSV heating and cooling that can be responsible for crack initiation. Figure 2.18 [28] illustrates how the shear and normal stresses during heating and cooling can ultimately lead to complete TSV delamination. The previously mentioned Cu pumping effect after the cooling stage, related to the BEOL more than Si, is also included in the illustration.

The thermo-mechanical behavior of a TSV has a direct impact on the FEOL as well. After processing, the TSV cools down to room temperature and pulls on the surrounding Si causing local stress. Figure 2.19 a) [34] points out the nature of the stress in Si. In the surrounding of the TSV the stress components consist of radial tension and tangential compression. Figure 2.19. b) [34] shows a SEM obtained cross section of the TSV and its immediate surrounding. FEOL devices, in this image FinFETS, are positioned within a couple of um from a TSV spanning 5um in diameter. Since Si is

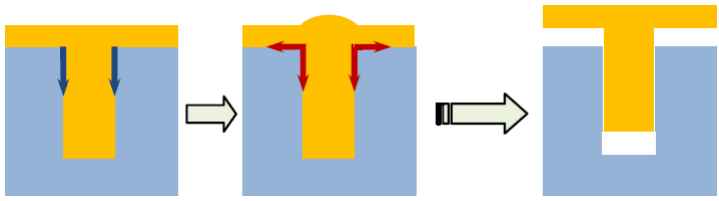


Figure 2.18: Illustration of shear and normal stresses during TSV heating and cooling, shown in figure 2.17, leading to complete TSV delamination from Si. [28]

a piezoresistive material (Chapter 3), the electrical performance of FEOL devices can change. Figures 2.20 a) and b) [34] present electrical measurements from planar transistors and FinFETs in the vicinity of a TSV. In this case, a TSV 5um in diameter and 50 um deep caused significant current shift in the FEOL devices ranging from several percent to above 10 % as the devices get closer to the TSV. Furthermore, a different impact level is visible when comparing device types and their relative position around the TSV. Close attention is paid today to the amount of stress the TSV generates as well as the sensitivity of different FEOL device types to stress. Evaluation of FEOL device sensitivities to mechanical stress is a part of this PhD work as well, with the difference that the study is further connected to other CPI effects as the TSV impact is primarily studied in other work.

2.2.1.2 3D IC stacking options

As discussed in section 2.1.2 the evolution of 3D stacking exhibited various technologies for vertical placement and interconnection of Si dies. In 3D IC stacking electrical connections are based on the positions of the TSVs. To this point 2 stacking approaches have been studied to a larger extent:

- Direct Cu-Cu bonding [27,36-40]
- Cu-Sn based microbumps [41-43]

The Cu-Cu bonding approach involves direct connection of TSVs over Cu landing pads. This provides a compact die stack with minimal space left in between dies and direct connection of the TSV with the BEOL of the other die. Figure 2.21 [35] illustrates a 3-die stack of direct Cu-Cu bonded dies. If a flat Cu bonding pad is used, the method is referred to as conventional Cu-Cu bonding. Figure 2.22 [39] gives an SEM image of a Cu-Cu bonding cross section. A 25 um Si die with 5 um in diameter TSVs is placed on Cu landing pads of the bottom tier.

An alternative approach to conventional bonding is insertion bonding. In this case, instead of a flat Cu landing pad, the TSV is inserted in a Cu plated cavity. Figure 2.23 a) [27] presents a finite element model of the insertion bonding method. Next to it, figure 2.23 b) [27] is a finite element model of the conventional Cu-Cu bonding with flat Cu landing pads. Figure 2.24 [27] presents an SEM image of a prototype 3D IC stack made with the Cu insertion method. The TSV fits within the Cu landing cavity and additionaly deforms by thermocompression to create an electrical connection.

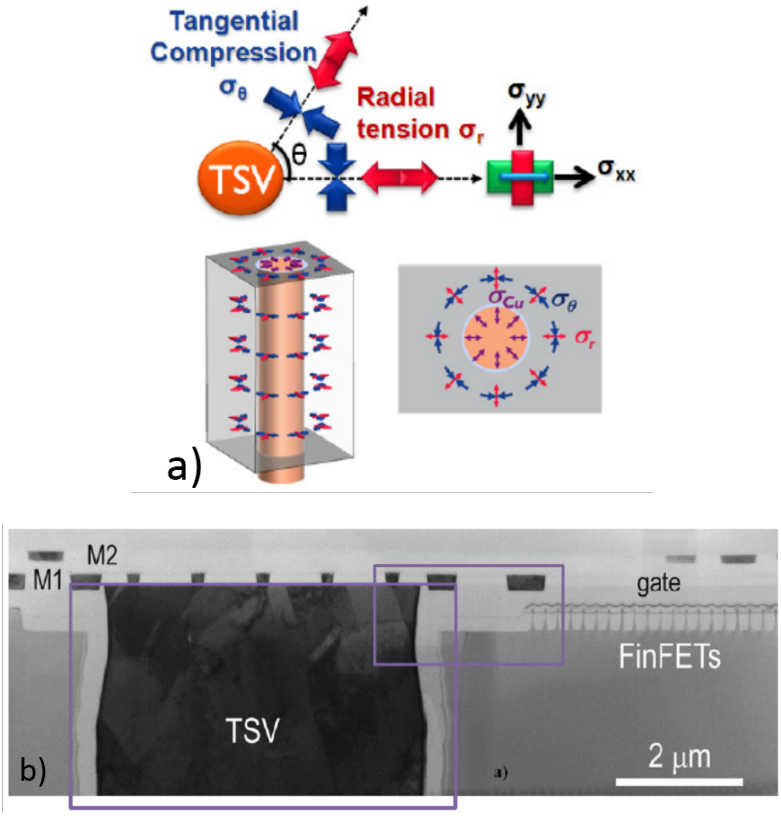


Figure 2.19: a) nature of stress in Si around a TSV, b) FEOL SEM image indicating proximity of FEOL devices to TSV [34]

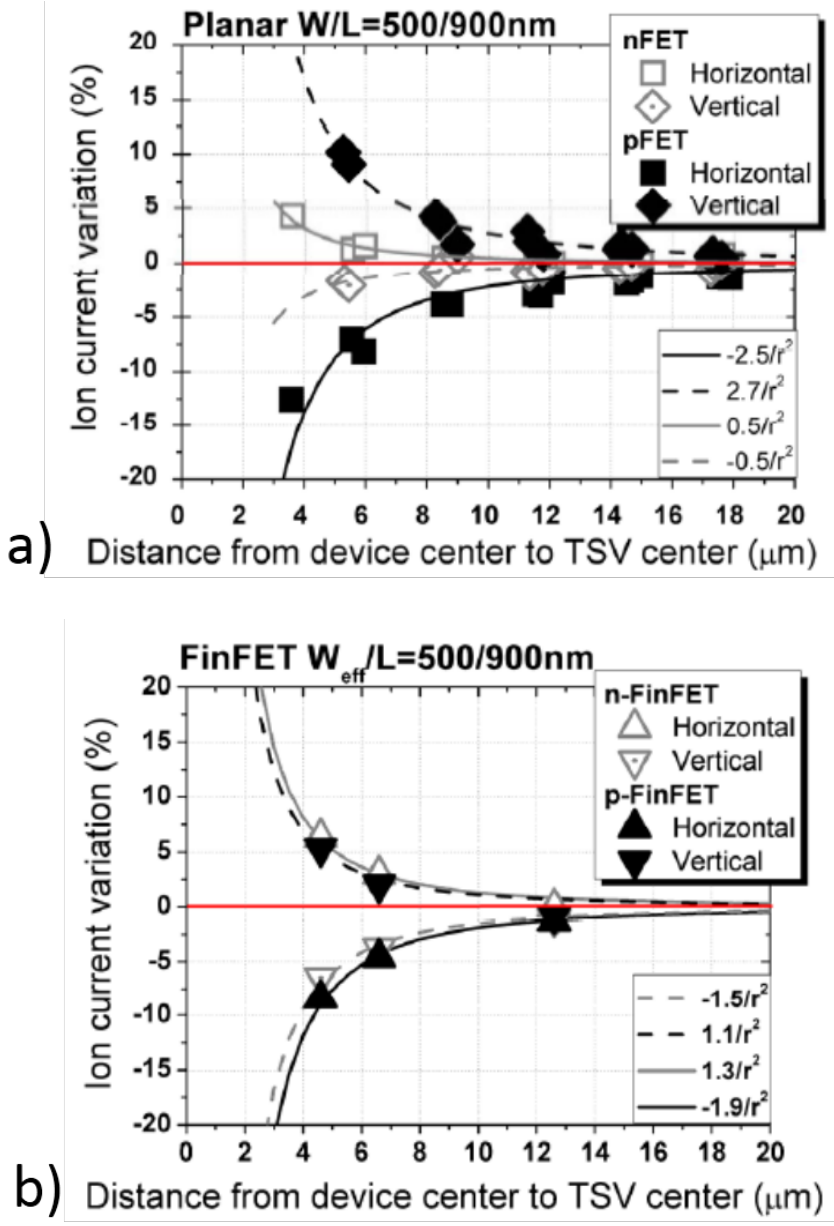


Figure 2.20: Current shifts observed on FEOL devices due to proximity to a TSV: a) n-type and p-type MOSFETs with current in [110] direction, horizontal, and current in [-110] direction, vertical, and b) n-type and p-type FinFETs with current in [110] direction, horizontal, and current in [-110] direction, vertical. Current shifts on MOSFETs and FinFETs are higher the closer they are placed to the TSV. [34]

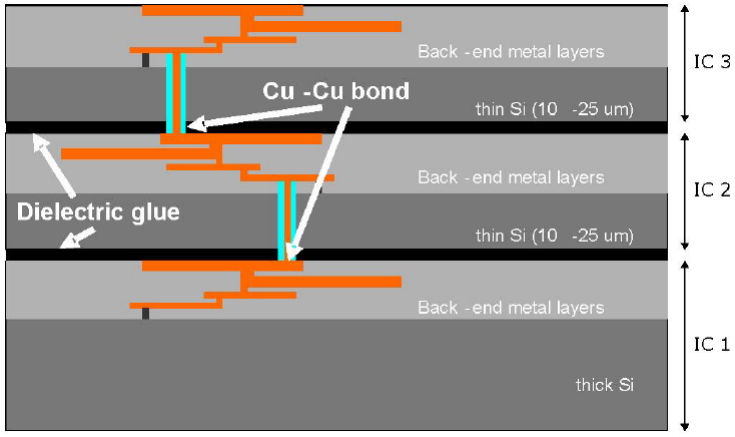


Figure 2.21: 3-die stack Cu-Cu bonding [35]

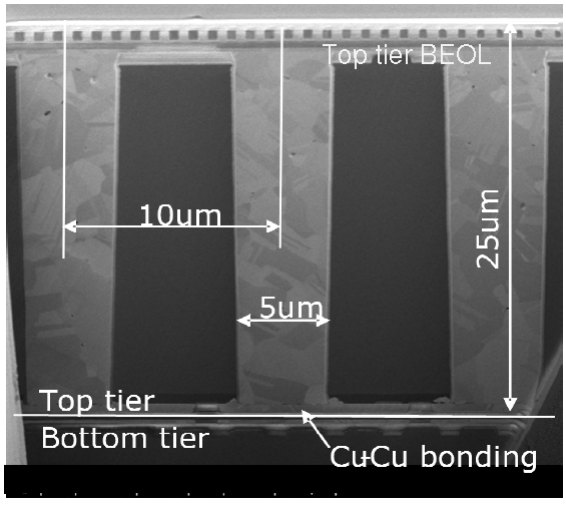


Figure 2.22: SEM image of a Cu-Cu bonding cross section [39]

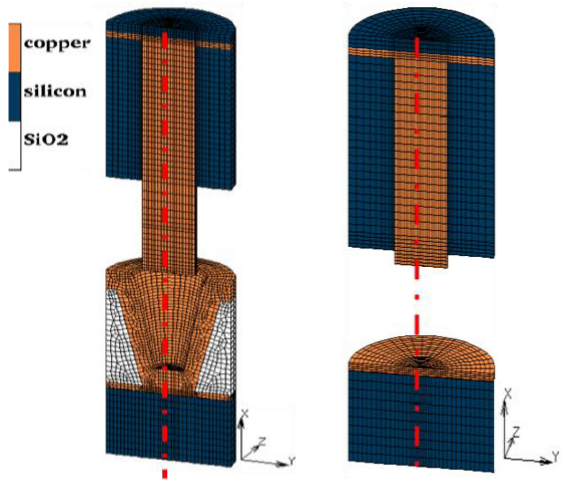


Figure 2.23: Finite element models indicating a) the basic principle of the Cu insertion method compared to b) the standard Cu-Cu bonding [27]

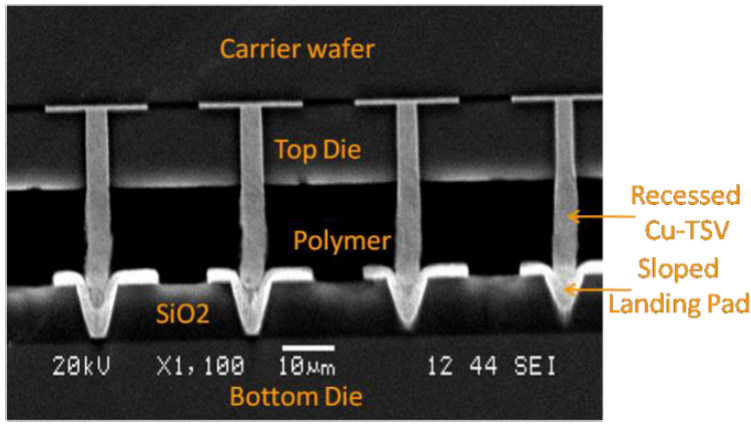


Figure 2.24: SEM image of a Cu insertion bonded stack [27]

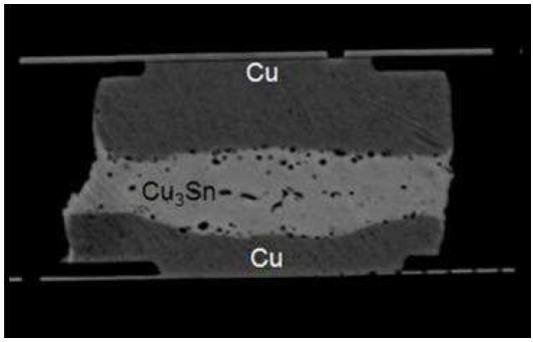


Figure 2.25: Microbump after Cu-Sn bonding [43]

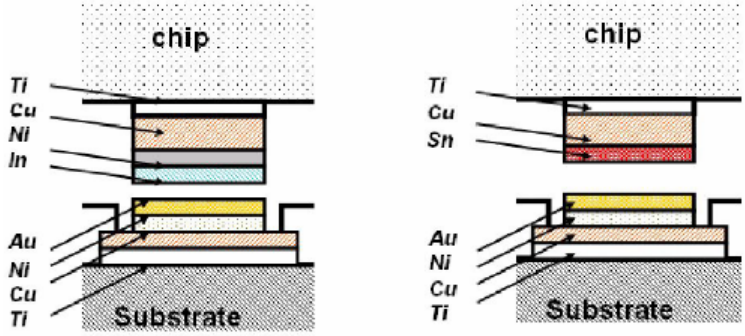


Figure 2.26: Advanced Cu-Sn microbump structures [41]

The second 3D IC stacking approach, that nowdays takes most of the attention involves usage of microbumps. Microbumps are usually sets of circular multilayer materials positioned between dies enabling vertical connection between dies. While the same can be said for solder balls connecting a microelectronic package and a PCB or for flip chip balls connecting dies with the laminate inside a BGA package, the microbumps are smaller within a range of tenths of micrometers in diameter. The 2 constituent materials are Cu and most often Sn. The microbump consists of 2 Cu pads, one on each Si side of the stacked dies, and Sn sandwiched in between the Cu. Figure 2.25 [43] shows a cross section SEM image of a microbump after bonding completion. Due to high temperature profiles during bonding the Cu and Sn react to create a brittle Cu-Sn intermetallic. The intermetallic chemistry depends on several factors, most prominent time of exposure on high temperatures. In general, initial Sn transformation stages are characterized by Cu_6Sn_5 material while the transformation is considered complete when all Sn is turned into Cu_3Sn. Additional thin material films can be added to this Cu-Sn structure such as Ni, as presented in figure 2.26 [41].

Microbumps are accompanied by usage of underfill material which fills the gaps between the Si dies around the microbumps acting as Cu-Sn surface cleaner and stress

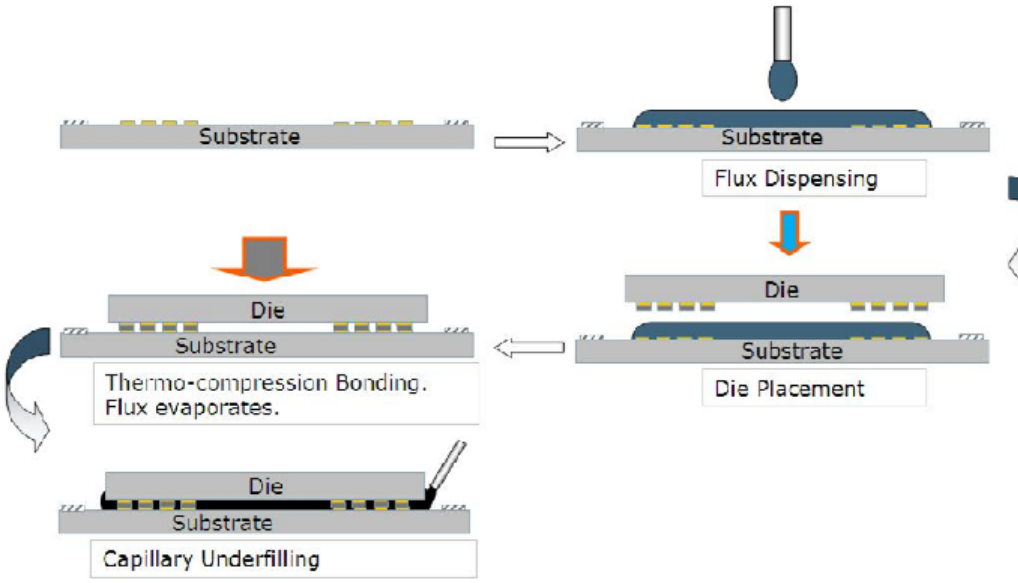


Figure 2.27: Capillary underfilling stack procedure [42]

buffer layer. More on stress related to 3D stacking in the following section.

There are several ways of incorporating underfill in the 3D IC stack. Two used in prototypes within this thesis are:

- capillary underfilling
- no-flow underfill (NUF)

The 2 procedures are presented in figures 2.27 and 2.28 [42], respectively. In the capillary underfilling process, the flux ensuring good wetting and cleaning of Cu-Sn surfaces and underfill placement and curing are done in separate processes. Firstly, flux is placed on the bonding side of the bottom Si die. The top die is subsequently aligned with the bottom one and pressure is applied until it reached a desired value. The structure is now heated to temperatures up to 250ºC to allow Cu-Sn bonding. After the process has finished and all flux used, underfill is dispensed in between the bonded dies as if it is flowing through small capillaries of the stack. The stack is then submitted to another heating stage for underfill curing.

The NUF is a flux filled polymer material that allows cleaning the bonding surfaces, Cu-Sn joining and underfill curing at the same time. In this process, the initial reflow process with flux is skipped and the underfill is placed immediately on the surface of the initial die to be bonded. The top die is aligned with the bottom one and a thermo-compression bonding with temperatures up to 250ºC occurs resulting in simultaneous Cu-Sn bonding and underfill curing.

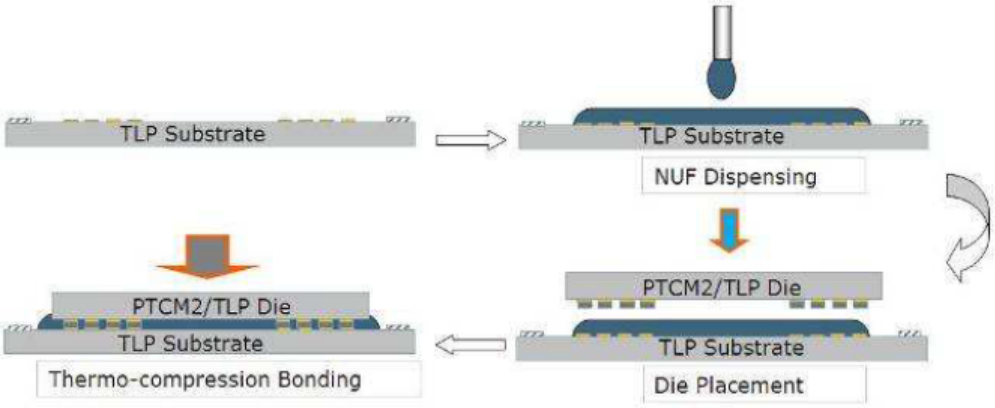


Figure 2.28: No-flow underfill stack procedure [42]

2.2.1.3 3D IC CPI

Chip-package interaction in 3D IC technology can in some capacity be annotated to the impact of the TSV as well (sections 2.2.1.1 and 2.2.1.2). However, its origin is the interaction of the Si die and all the constituents of a package surrounding that Si die. According to that definition, impacts are revealed after assembly steps subsequent to die processing meaning susbsequent to FEOL, TSV and BEOL processing.

Along with the introduction of 3D IC technology, as one of the methods to decrease signal delay and power dissipation, low permittivity dielectrics, referred to as low-k materials, are being introduced in the BEOL. As much as these materials may enhance the electrical performance of interconnects, they are mechanically very fragile. The concern rises with their implementation in 3D stacks which can provide additional mechanical loads to the BEOL compared to a 2D configuration, during additional 3D assembly steps. The risk and analysis of implementing fragile Cu/low-k BEOL in 3D IC stacks is being widely investigated [44-52] and is usually the first topic that occurs in literature concerning CPI in 3D ICs. Stress generation, crack propagation and in-situ failure analysis are investigated in finite element models and processed BEOL prototypes. Figure 2.29 [48] shows an image of a 10-layer low-k BEOL where a typical crack is observed, in this case fluctuating between layers M7 and M8.

Finite element models are employed to understand the occuring crack inititation and propagation linked to low-k material. Figure 2.30 [48] presents an example of a developed finite element model consisting of 4 submodels, starting from stack level down to BEOL level. Submodels are utilized to tackle variations in dimensions from package level to BEOL interconnect level and ensure fine meshing density on each level. Running several submodels of a structure rather than one big model cuts down on time and required simulation resources. This modeling approach is often refered to as global-local modeling. Figure 2.31 [48] reveals the results of the finite element model where the highest chance for crack propagation, based on obtained nodal release energy (NRE) values, is observed in metal layers M7 and M8, consistent with findings

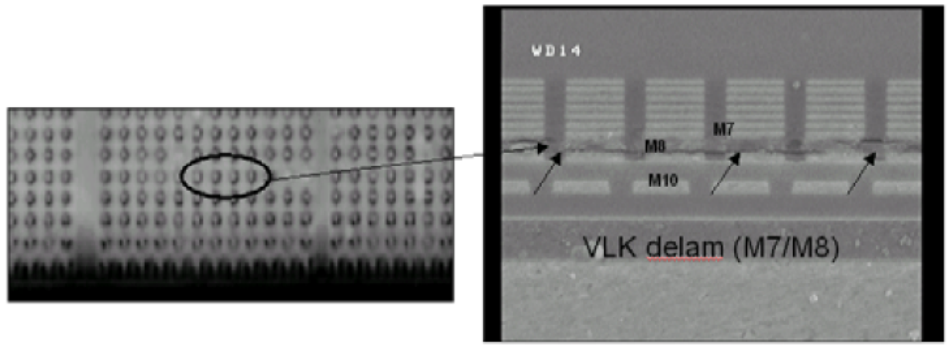


Figure 2.29: Crack observed between the 7th and 8th metal layer in a 10-layer BEOL [48]

from figure 2.29 [48]. NRE is related to fracture mechanics and is a measure of energy. The nodal release energy is proportional to the energy release rate. The energy release rate is the energy dissipated during fracture per unit of newly created fracture surface area. The energy applied to the crack tip needed to initiate crack propagation is equal to the dissipated energy during crack propagation. Crack propagation tends to occur at the interface having a higher NRE value.

When it comes to 3D IC stacking, studies are mostly dedicated to the fundamentals of stacking technology [53,54] such as thin wafer handling and alignment or research is performed on the chemistry of microbumps [43,55-57]. Figure 2.32 [54] shows a cross section of a 2-die stack joined with Cu-Sn based microbumps where the stacking process was being optimized for fine pitch small microbumps, in this case 8um microbumps with a pitch of 15um.

Furthermore, attention is being paid to the electrical functionality of the 3D IC system in terms of designing system arhitecture and circuitry for 3D, determining signal propagation and testing particular circuit performances [58-60]. Closely related, system level 3D IC test methodologies and accessibility of structures are discussed [61,62]. Figure 2.33 and 2.34 present a study from [59] focusing on electrical modeling of a TSV and microbump and signal propagation through them. Figure 2.33 a) presents the basic schematic with the driver and load while figure 2.33 b) the expanded electrical model for the TSV and microbump. Signal delay through the TSV-microbump structure versus increasing capacitive load C_{ESD} is presented in figure 2.34. In other studies application of system architectures such as processor-memory interaction are explored. Figure 2.35 [58] presents a microcontroler arhitecture developed for 3D IC technology with several stacked DRAM layers on top of a processor layer.

To meet customer's product-quality expectations, each individual IC needs to be tested for manufacturing defects that might have incurred during many processing and assembly steps in 3D IC [61]. These tests should be both effective and cost-efficient. Figure 2.36 presents a testing concept flow discussed in [61] for wafer-level and package-level tests, with respect to test contents and wafer-level probe access, and the on-chip infrastructure required for 3D ICs.

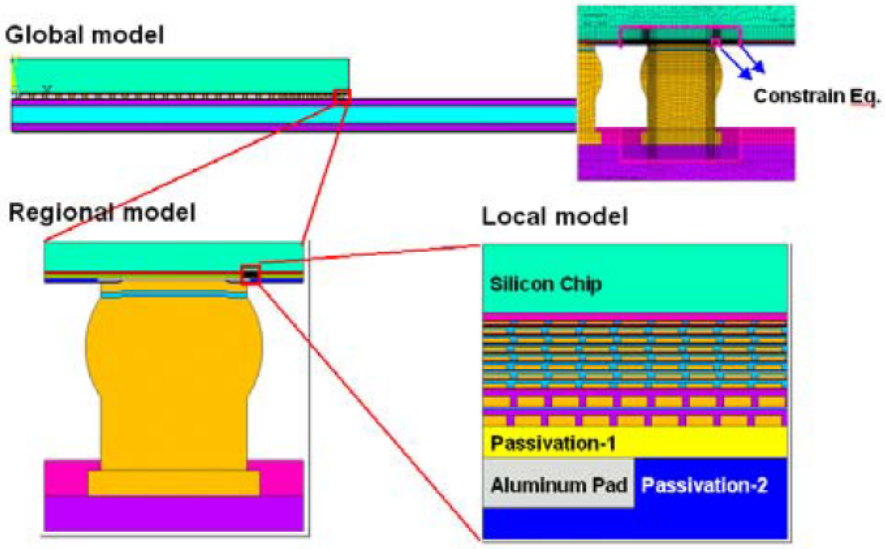


Figure 2.30: Finite element model built to investigate BEOL cracking [48]

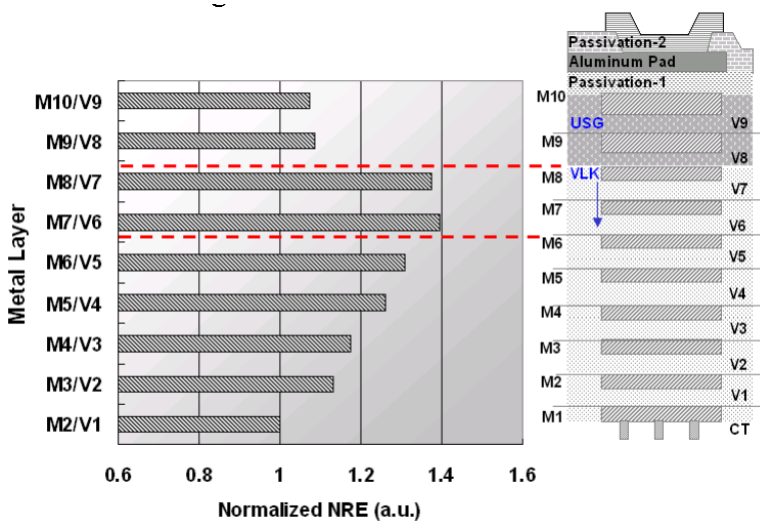


Figure 2.31: Result of the finite element model indicating highest nodal release energy (NRE) in the area of the 7th and 8th metal layers [48]

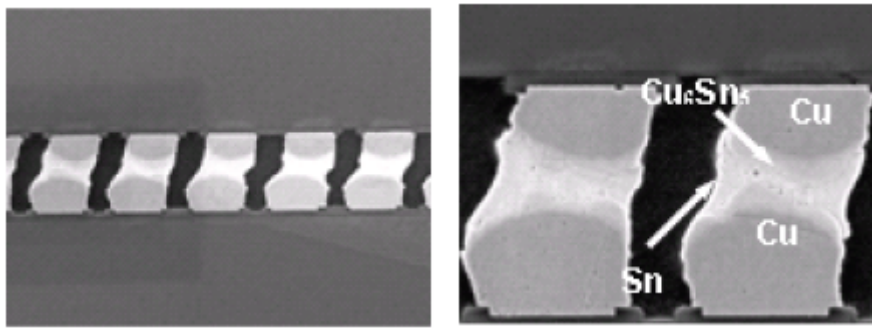


Figure 2.32: Optimized fine pitch Cu-Sn microbumps [54]

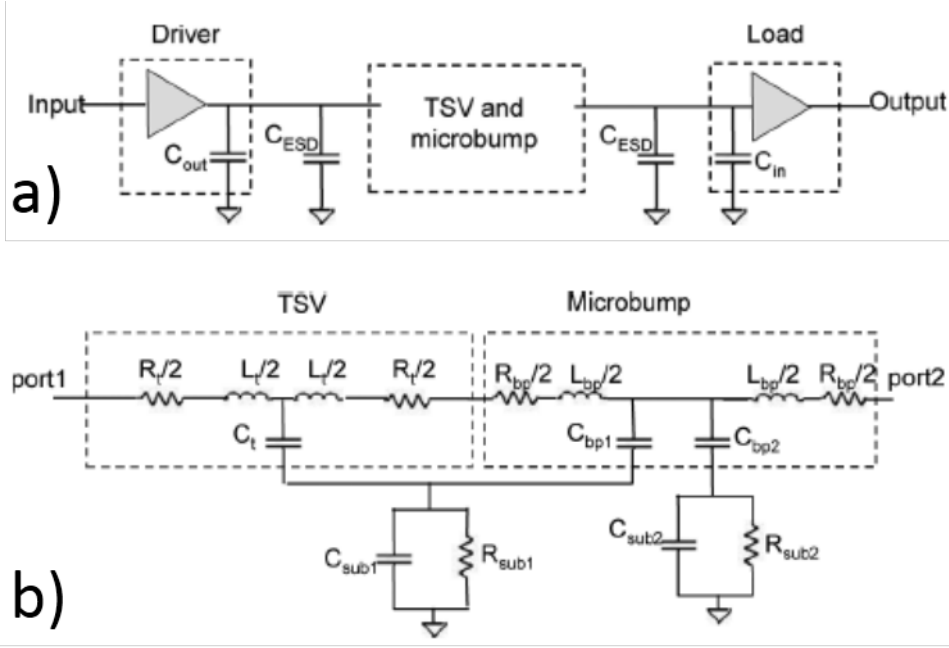


Figure 2.33: Modeling signal propagation through a TSV and microbump: a) the block schematic with inputs and outputs, b) the TSV and microbump equivalent circuit [59]

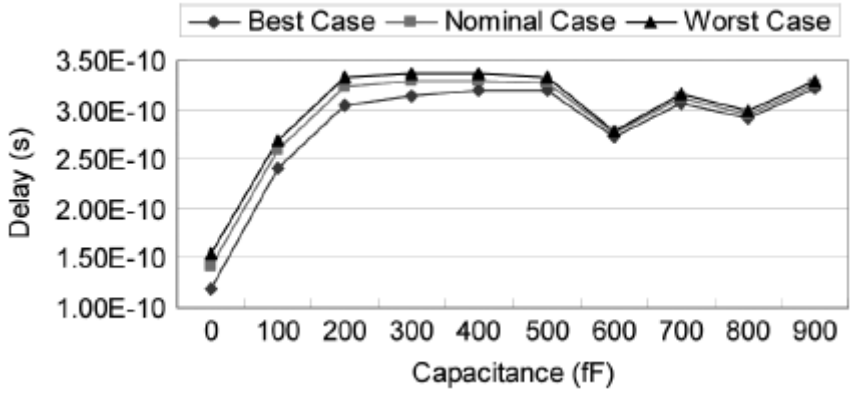


Figure 2.34: Resulting delays from several TSV and microbump modeled cases [59]

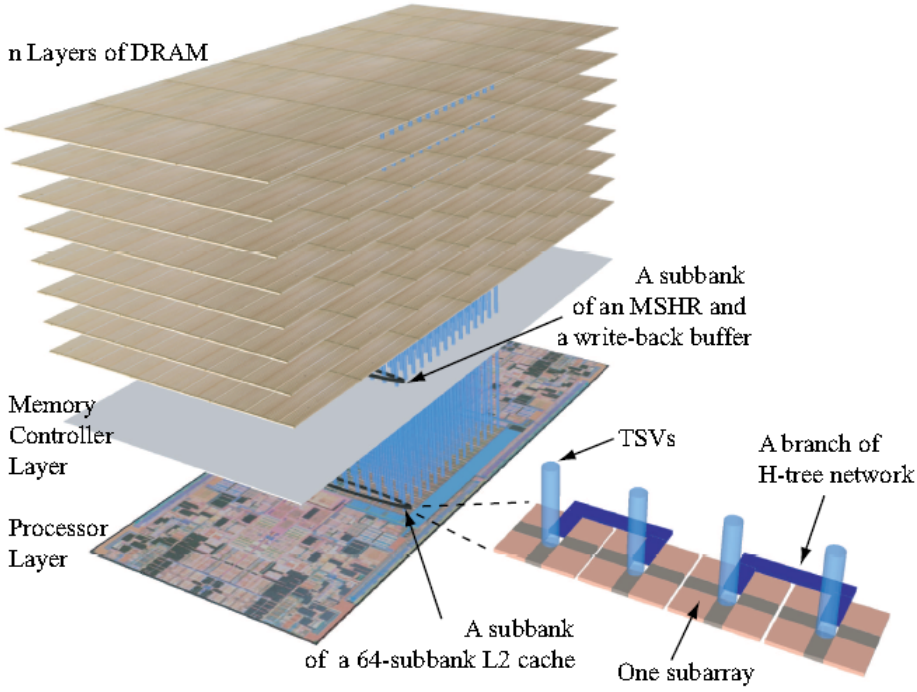


Figure 2.35: Example of a developed 3D arhitecture for conventional 2D circuits - several DRAM layers on top of a processor die [58]

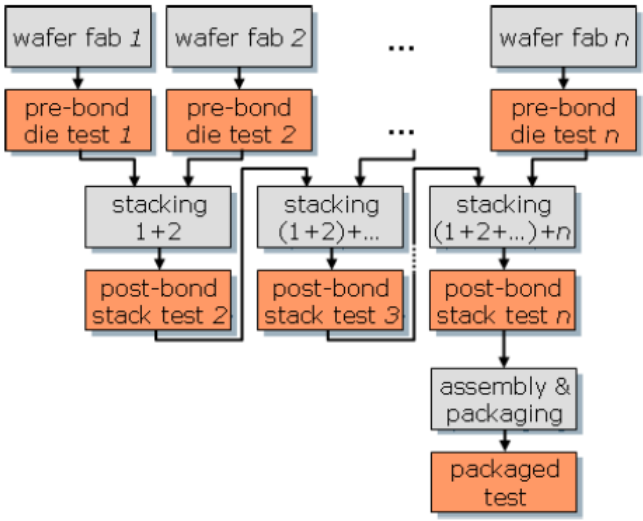


Figure 2.36: Testing concept flow for wafer level and package level tests [61]

The CPI review in this paragraph until now showed presence of studies related to 3D IC electrical properties, architecture concepts, materials and testability. Thermal problems were not directly addressed as they are out of scope of this thesis. However, thermal properties of 3D IC circuits are also one of the bottlenecks of this technology and are a part of CPI. More information on thermal research in 3D ICs can be found here [63-68]. A crucial component of CPI and for this thesis is thermo-mechanical interaction. In publications prior to this thesis, thermo-mechanical CPI studies in 3D IC were directed either directly to the BEOL [44-52], indirectly to UBM and BEOL failures as a consequence of microbump impact [48,73], stress in the microbumps [55], or monitoring of global thin die and stack warpages [71]. Review papers including thermo-mechanical impact in 3D IC [69,70] state these topics, along with TSV impact, as main concerns as well. Figure 2.37 taken from [71] points out a study focused on warpage of a 50 μm thin die. The underfilling process can cause excessive warpage that can cause inability of further stacking or packaging and increase TSV stress in Si. Figure 2.37 a) gives the basic cross section schematic of the explored thin die on substrate, while figure 2.37 b) presents an SEM image of the cross section with visible warpage of the thin die. Figure 2.38 [72] presents a finite element model simulating mechanical stress in a 4-layer 3D IC stack after dissipation of 0.25W on the bottom side of each Si die. The 3D IC stack is assembled to a substrate and PCB. The points of interest in the study are presented in further detail in figure 2.39 a) - c) [72]. Figure 2.39 b) indicates stress distribution in the solder balls, figure 2.39 c) indicates the stress in TSVs and microbumps while 2.39 a) gives an overview on global stress in the dies. The stress in the dies is mentioned to be uniform apart from the areas in the vicinity of TSVs. The origin of the stress in Si is not elaborated or investigated in more detail apart from linking the stress to TSV presence.

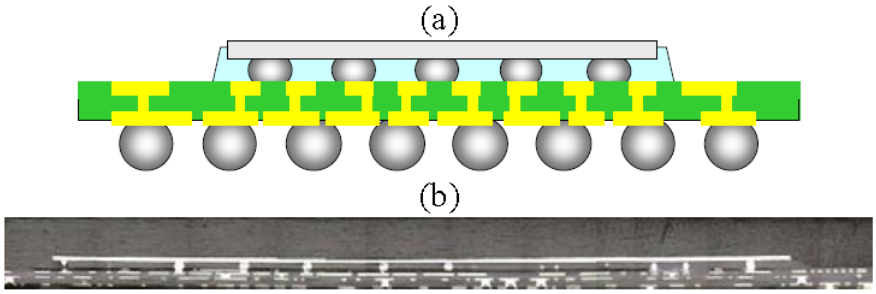


Figure 2.37: Investigation of warpage of a thin Si die: a) Illustration of a thin Si attached to a substrate and b) its realized counterpart [71]

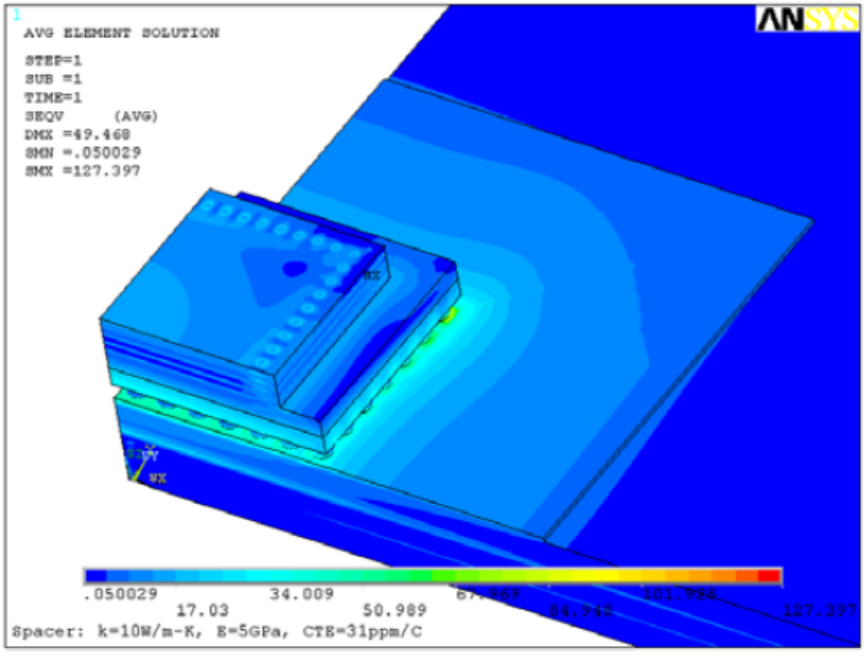


Figure 2.38: Finite element model of a 4-layer 3D IC stack assembled on a substrate and PCB. Stress is monitored after dissipation of 0.25 W on the bottom of each Si die. [72]

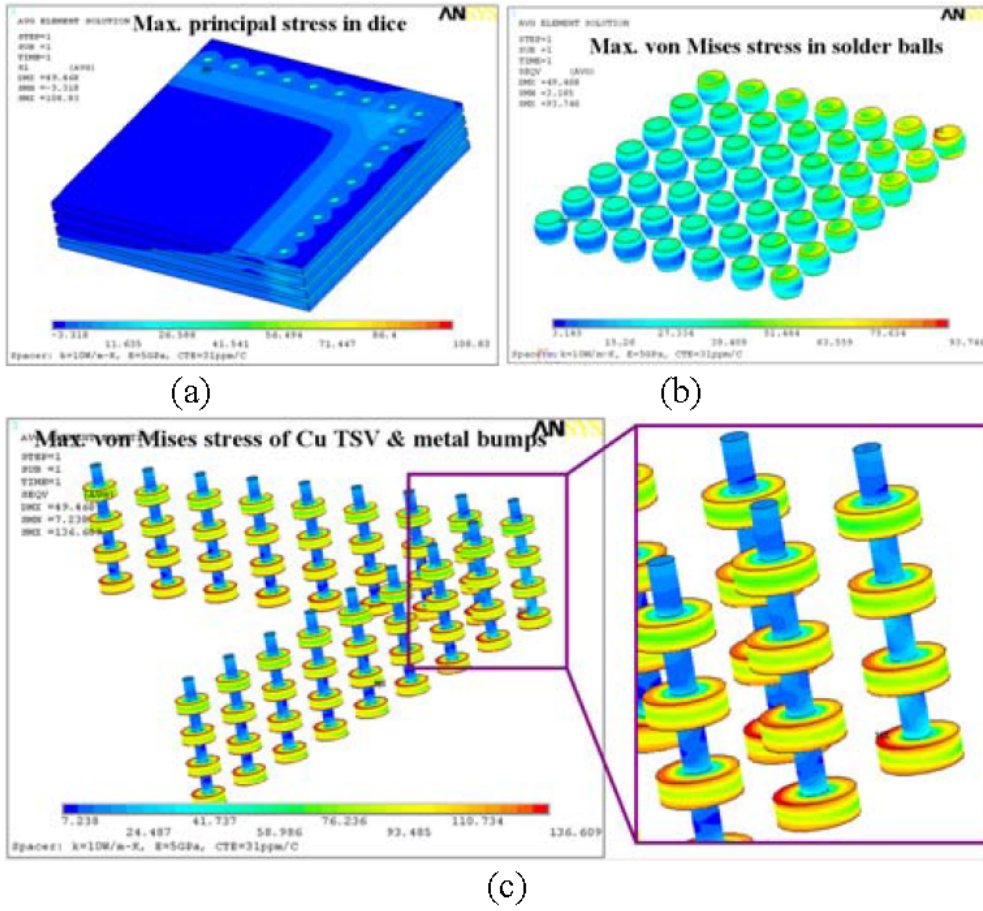


Figure 2.39: Stress points of interest of the finite element model after dissipation of 0.25 W on the bottom of each Si die: stress in a) dies on a general scale, b) solder balls and c) microbumps. Local impact on FEOL is not investigated. [72]

The thermo-mechanical impact on the FEOL was limited to the impact of TSVs [22, 32-34, 69,70]. To the start of this PhD in July 2010, no CPI impact on the FEOL within 3D SIC with origin other than TSV impact was explored.

2.2.2 Stress sensors

Si has a distinctive electrically measurable reaction to mechanical stress and many Si based devices within microelectronics have utlilized this property. Metal based sensors were also explored [74-80] however their implementation in the semiconductor industry was limited due to their temperature instability.

Figure 2.40 categorizes piezoeffects in Si into 5 groups [81]:

- Piezojunction
- Piezotunneling
- PiezoHall
- Piezoresistive
- PiezoMOS

The piezojunction effect refers to the changes of the saturation current of a bipolar transistor or a p-n junction under mechanical stress. It was revelead in 1951 and investigated most thoroughly in the 1960s when it was found that the effect is large for high, anisotropic stresses. The study of the piezojunction effect and its consequences for circuits and sensors is the main topic of [81].

The essence of the piezotunneling effect lies in the reverse-biased heavily doped, shallow, lateral junction. The reverse current flowing through the lateral junction is dominated by band to band tunneling and since strain in Si affects the energy and shape of the bands, the tunneling depends on the stress. The temperature dependence observed with the piezotunneling effect is an order of magnitude lower than with the piezoresistance effect, however the sensitivity to stress is about 4 times lower when compared to the piezoresistance effect.

The Hall effect is known to describe the impact of a magnetic field on electric current. The Hall sensors use this effect to their advantage as the current sensitivity to magnetic fields is altered under mechanical load. Hall stress sensors have a low-temperature coefficient.

The piezoresistive effect is the most commonly utilized effect in solid-state sensors, particularly in microelectronics as Si, as the most common material in the industry, is fundamentally a piezoresistive material. The basis of the effect is the change in material resistivity due to impact of mechanical stress. The piezoresistive effect and piezoresistive sensors are intensively used within this thesis. Section 3.1 elaborates on the piezoresistive effect in more detail while this section focuses on piezoresistive stress sensors that were previously used in a certain form to measure the impact of mechanical stress in ICs.

Regarding piezoresistive device types, Si resistors are the most explored ones due to processing and integration simplicity. Resistor rosettes in various configurations

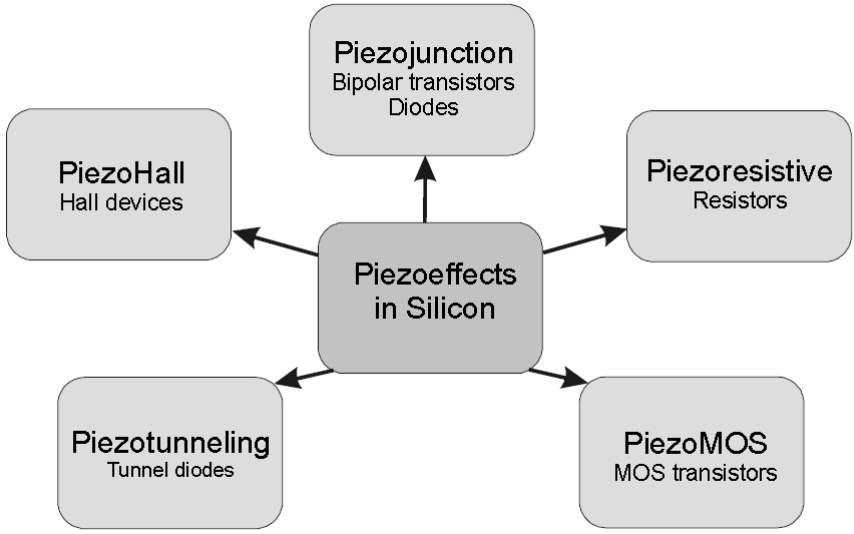


Figure 2.40: Piezoeffects in Si according to [81]

were proposed in an attempt to capture all stress components in Si [82,83]. Figure 2.41 [83] presents a 4-element rosette on (100) Si surface, pointing in 4 different Si crystal orientations. Equation 2.1 [83] presents the temperature compensated stresses that are obtainable with this configuration, a difference of two in-plane stresses and one shear stress. Temperature has a direct effect on stress extraction as the piezocoefficient values π connecting stress to resistivity shift are highly dependent on temperature fluxuations. Measurement errors can be a consequence of several factors such as:

- Instability of temperature during measurement or variable temperature between two measurements especially if measurements are repeated after longer periods of time
- Calibration drift of measurement instruments
- Inaccuracy of temperature measurements

In equation 2.2 [83], the two in-plane stresses, $\sigma_{11}^{'}$ and $\sigma_{22}^{'}$ can be individually extracted without temperature compensation, hence the addition of terms $\alpha_1 T$, where α represents the temperature coefficient and T the temperature. With respect to the [100]/[010]/[001] reference frame π_{44}^{p} represents the p-type shear piezocoefficient, π_S stands for the sum of the p-type in-plane piezocoefficients $\pi_{11}^{p} + \pi_{22}^{p}$ and π_D for the difference of the p-type in-plane piezocoefficients, $\pi_{11}^{p} + \pi_{22}^{p}$.

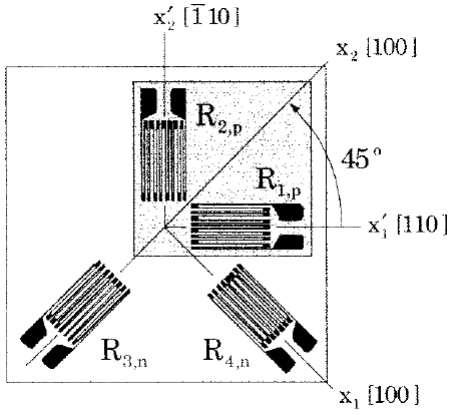


Figure 2.41: Standard 4-element resistor rosetter as a stress sensor [83]

$$(\sigma'_{11} - \sigma'_{22}) = \frac{1}{\pi^P_{44}}\left[\frac{\Delta R_1}{R_1} - \frac{\Delta R_2}{R_2}\right]$$
$$\sigma'_{12} = \frac{1}{2\pi^n_D}\left[\frac{\Delta R_3}{R_3} - \frac{\Delta R_4}{R_4}\right] \tag{2.1}$$

$$\sigma'_{11} = \frac{\left[\frac{\Delta R_1}{R_1} + \frac{\Delta R_2}{R_2}\right]}{2\pi_S} + \frac{\left[\frac{\Delta R_1}{R_1} - \frac{\Delta R_2}{R_2}\right]}{2\pi_{44}} - \frac{\alpha_1 T}{\pi_S}$$
$$\sigma'_{22} = \frac{\left[\frac{\Delta R_1}{R_1} + \frac{\Delta R_2}{R_2}\right]}{2\pi_S} - \frac{\left[\frac{\Delta R_1}{R_1} - \frac{\Delta R_2}{R_2}\right]}{2\pi_{44}} - \frac{\alpha_1 T}{\pi_S} \tag{2.2}$$

The same paper extrapolates a very useful general formula relating resistance shift in Si for resistors in an arbitrary orientation on (100) Si wafers. In equation 2.3, the relative resistance shift $\frac{\Delta R}{R}$ is related to normal stress components σ'_{11}, σ'_{22}, σ'_{33} and one shear component σ'_{12} via combinations of piezocoefficients π and the angle of rotation on the Si surface, ϕ. $\phi = 0$ implies the resistor placed on a (100) Si surface is pointing in the [110] direction, as indicated on the accompanying figure in 2.42 [83]. Apart from stress, resistance is sensitive to change of temperature as well, hence the αT terms at the end of the equation. The impact of temperature on resistance shift is not negligible.

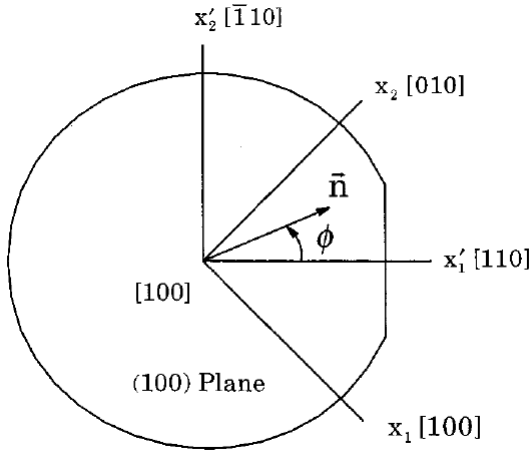


Figure 2.42: Reference frame for equation 2.3 [83]

$$\begin{aligned}\frac{\Delta R}{R} = &\left[\left(\frac{\pi_{11}+\pi_{12}+\pi_{44}}{2}\right)\sigma'_{11}\right.\\ &\left.+\left(\frac{\pi_{11}+\pi_{12}-\pi_{44}}{2}\right)\sigma'_{22}\right]\cos^2\phi\\ &+\left[\left(\frac{\pi_{11}+\pi_{12}-\pi_{44}}{2}\right)\sigma'_{11}\right.\\ &\left.+\left(\frac{\pi_{11}+\pi_{12}+\pi_{44}}{2}\right)\sigma'_{22}\right]\sin^2\phi\\ &+\pi_{12}\sigma'_{33}+(\pi_{11}-\pi_{12})\sigma'_{12}\sin 2\phi\\ &+\left[\alpha_1 T+\alpha_2 T^2+\cdots\right]\end{aligned} \tag{2.3}$$

Other resistor based devices were also proposed, such as the Van der Pauw sensor [84]. In essence, the Van der Pauw sensor in its most simple form is a square shaped 4-terminal resistor. Figure 2.43 a) and b) present some of the possible configurations of the Van der Pauw sensor. The Van der Pauw sensors provide similar stress component extraction capability in a more compact and scalable form [84].

The MOSFET is the most commonly used device in microelectronics and its sensitivity to stress was initially investigated already in the 60s [81]. The mechanical stress modifies the mobility of the majority carriers resulting in drain current shift. MOSFETs have not been often analyzed as stress sensors, still compelling literature on their potential is available. Analysis has been done on individual MOSFETs [85,86] and

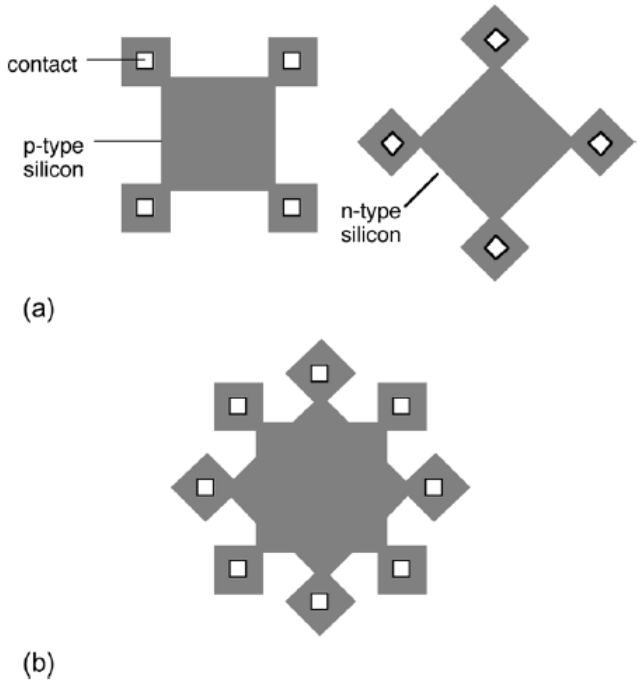


Figure 2.43: Van der Pauw stress sensor configurations: a) p-type and 45 degrees rotated n-type and b) functionality of two separate sensors combined in one [84]

development of MOSFET based sensor circuitry [87-89]. Since doping has a large impact on piezocoefficients (see section 3.1), MOSFETs have the potential to be more sensitive than resistors due to their lightly doped channel region compared to often heavily doped Si resistors [88]. Figure 2.44 a) and b) [85] present PMOS and NMOS drain current sensitivity to stress results from 500 um channel transistors with current flow in [110] direction. Several tens of MPa are reflected in a detectable several percent of drain current shift. The advantage of producing MOSFET sensor circuits rather than utilizing individual components is temperature compensation leading to temperature compensated differential of in-plane stress components and an individual shear component, similar to equation 2.1. Figure 2.45 [87] presents MOSFET circuitry proposals for extracting the difference in in-plane stress components, figure 2.45 a) and for extracting one shear stress component, figure 2.45 b).

Most recent studies on MOSFETs as stress sensors involve new conceptual designs with devices unofficially labeled Pseudo-Hall transistors [90,91] due to their resemblence to the Hall effect. Magnetic fields are however not involved in their operation. Although they are also able to provide temperature compensated differences of in-plane stress sensors and a shear stress component, they do so with integration in one transistor, ommiting the need for a MOSFET circuit as in figure 2.45.

When mechanical stress is applied to MOSFETs, splitting, distortion, and shifting of the energy band structure occurs [90]. The mobility tensor of the transistor exhibits off-diagonal terms resulting in a pseudo-Hall electric field perpendicular to the direction of the drain to source current. This occuring potential can be measured as a stress dependent value. Figure 2.46 [90] illustrates the operating principle. Presuming the transistor channel with width W and length L is created, drain to source voltage V_{DS} results in current I_{DS}. The application of stress to the channel causes an electric field perpendicular to I_{DS} which is measured as a potential drop, V_{pH}. Figure 2.47 [90] depicts the usual appearance of the pseudo-Hall sensor. The usually square shaped channel is covered with a gate plate and contacted with four terminals in its corners, two for the drain and source contact and the other two for the pseudo-Hall contacts for V_{pH} measurements.

Other stress sensors resembling pseudo-Hall sensors have been developed, such as the one shown in figure 2.48 [92]. Although similar in appearance to the pseudo-Hall sensor, this piezo-FET does not base its operation on the transverse electric field but combines 4 separate MOSFETs in a Wheatstone bridge configuration within in one sensor. The stress component outputs are the same as with the pseudo-Hall sensor, temperature compensated difference in in-plane stresses and one shear stress component. However, the addditional advantage of this piezo FET is that it is not sensitive to outer magnetic fields.

In an attempt to extract individual stress components rather than differences in stress components, this thesis incorporates usage of standard individual MOSFETs as stress sensors having in mind that the surrounding in which they will be used, cleanroom or measurement lab, are under strict temperature control.

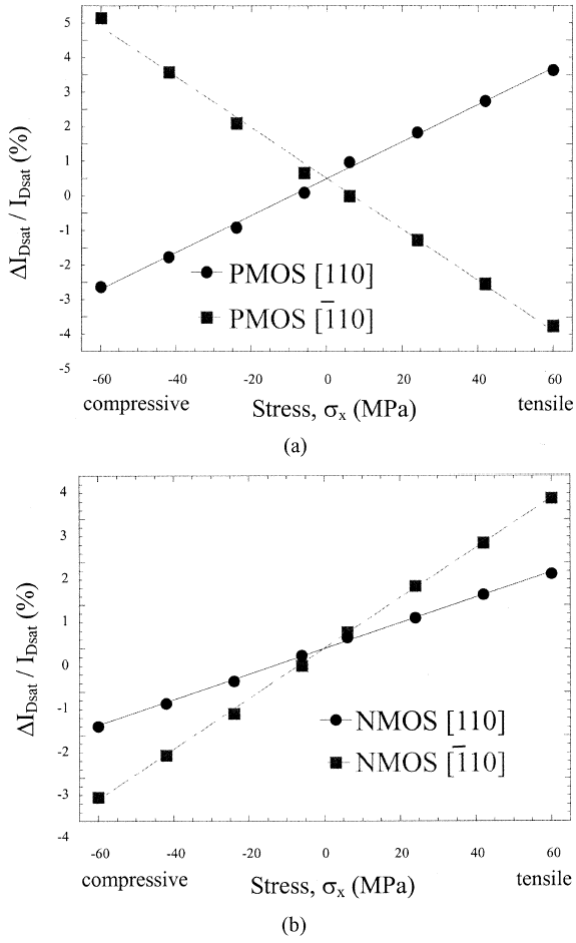


Figure 2.44: Electrical response of MOSFETs to mechanical stress: current shift observed with a) p-type and b) n-type MOSFET [85]

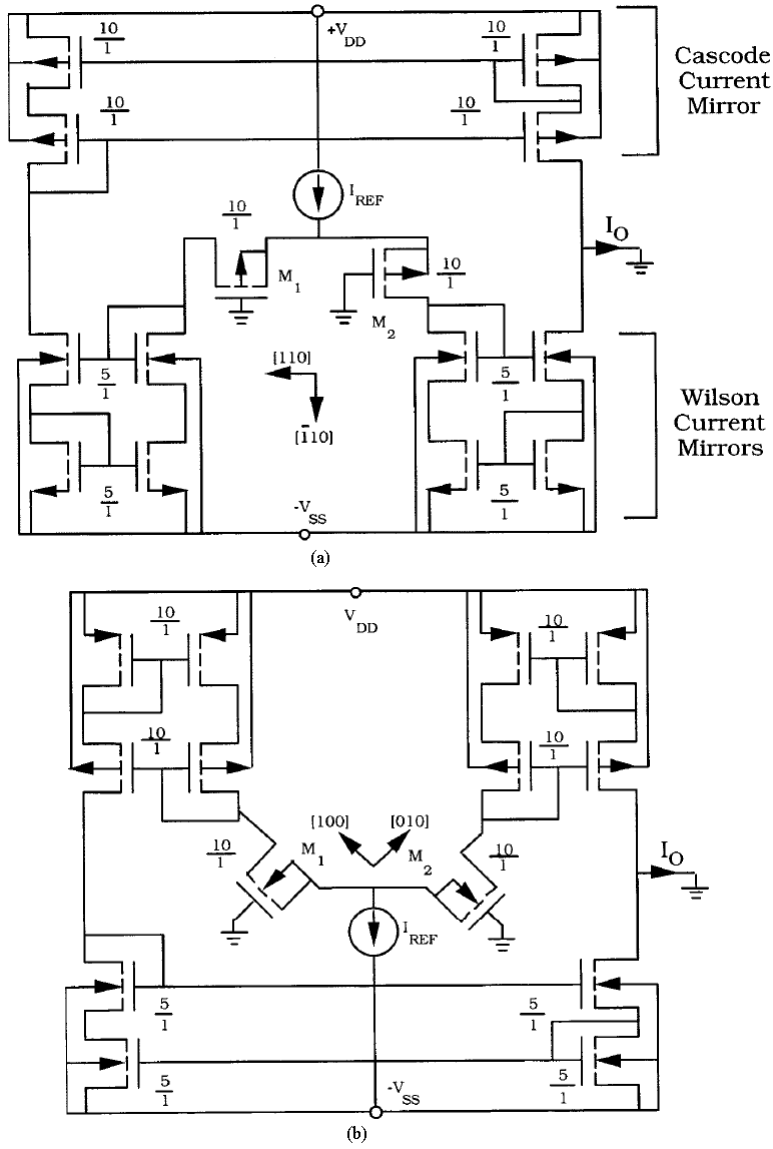


Figure 2.45: MOSFET circuitry developed to extract temperature compensated a) difference in in-plane stress components and b) one shear stress component [87]

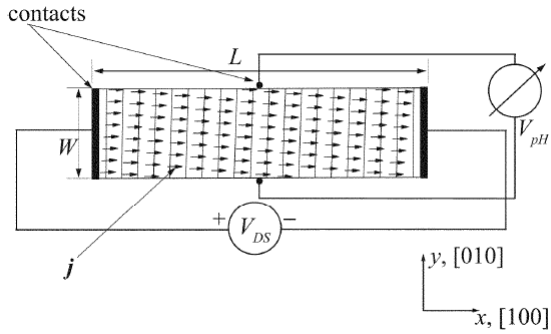


Figure 2.46: Operating principle of a Pseudo Hall transistor, measurement of V_{pH} perpendicular to current direction in conditions of applied mechanical load [90]

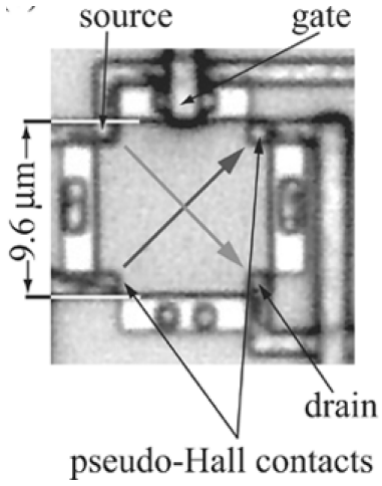


Figure 2.47: Stamdard appearance of a Pseudo Hall sensor, a square like gate covered transistor with 4 terminals in its corners [90]

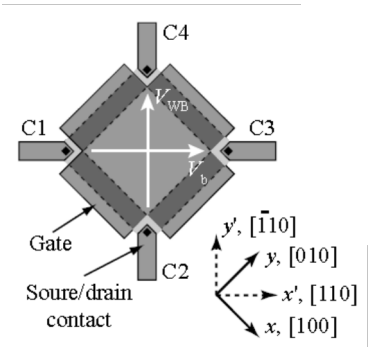


Figure 2.48: Similar to appearance but with a different operating principle, 4 MOSFETs in a Wheatstone bridge configuration combined into one sensor [92]

2.3 Contribution of this thesis

With respect to the historical and literature overview in section 2.1 and 2.2, a set of goals is given with which this thesis aims to continue to improve 3D SIC technology and leave a mark in the scientific community:

- Study the impact of the underfill-microbump stress mechanism on the FEOL in 3D SICs, using primarily FEOL stress sensors and finite element modeling
- Propose underfill-microbump stress mitigation guidelines concerning its impact on the FEOL
- Propose an evaluation methodology for stress sensors for 3D SIC CPI
- Evaluate and implement stress sensors in 3D SIC stacks and 3D SIC packages in order to enable monitoring local and global stress trends occuring during 3D SIC stack and package assembly

2.4 Summary

The idea of vertically stacking electrical circuits appeared already in the 1950s. Micromodules, stacked and packaged electrical circuits came to vast commercial production and proved to have soaring reliability. Micromodule production was affected by the invention of the integrated circuit in 1959. Although projected in 1962 that 3D electrical circuit stacks and integrated circuits could go on simultaneously in production and eventually meet to combine in one technology, micromodules were soon abandoned, integrated circuits took all the attention. Consistent industrial interest for 3D stacking of integrated circuits appeared again decades later, in the late 1990s and beginning of 2000s. Multi-chip modules (MCM) and systems-in-package (SiP) appeared first, where vertical placement of dies was observed on module level and lower level of integration package level, respectively. Wafer level packaging acted as an improvement

of SiP technology with package level vias connecting vertical levels consisting of one or more Si dies. All die-to-die interconnects were made off-die.

Monolithic 3D IC circuits, with stacking integration on FEOL and BEOL level within one die, appeared as an idea already in the 1980s but have mostly remained on idea level due to limitations of processing technology. 3D stacked integrated circuits (3D SIC) started developing with first studies on through-Si vias (TSVs) in the 1990s. A TSV enables direct vertical connection of dies within the die integrated circuit. Research and prototyping of 3D SICs launched on larger scale and exhibited rising industrial interest from 2004 coinciding with a TSV based 3D SIC prototype published by Intel.

3D SIC technology compared to further development of 2D ICs promises shorter interconnects, smaller PCB footprints and easier heterogenous technology integration within one stack and package. Shorter interconnects increase the performance of the system by decreasing signal delay and lowering power consumption.

Stacking options that have been explored until today include direct Cu-Cu bonding including Cu insertion bonding and bonding with Cu-Sn based microbumps. The most common process for vertically joining stacks with Cu-Sn based microbumps is thermo-compression bonding. The underfill material can be dispensed after thermo-compression, with capillary dispensed underfill, or prior to thermo-compression, with no-flow underfill.

Today, from a thermo-mechanical viewpoint, particular chip-package interaction (CPI) impacts are being investigated. Studies on reliability of microbumps are consistently published. On die level, most attention is being paid to BEOL mechanical integrity after package level assembly due to the increasing mechanical fragility of the BEOL with introduction of low-k materials. Studies covering thermo-mechanical impact on the FEOL are limited to the stress impact of TSVs. To the start of this PhD in July 2010, no CPI impact on the FEOL within 3D IC, with origin other than TSV impact, was explored. In relation to FEOL impact, this PhD reveals the impact of the underfill-microbump stress mechanism on the FEOL, monitors its evolution on several generations of 3D IC stacks and provides guidelines for its mitigation.

Several types of devices were used as Si stress sensors in microelectronics. They can be grouped in one of 5 groups, depending on the Si effect they are based on: piezoresistive sensors, piezo-MOS sensors, piezotunneling sensors, piezoHall sensors and piezojunction sensors. Piezoresistive and piezo-MOS sensors are explored within this thesis. Metal based piezosensors and in particular piezoresistive Si resistor rosettes have been previously reported. Metal piezoresistor properties are highly dependent on temperature. This PhD focuses on exploitation of FEOL active devices, primarily MOSFETs as Si stress sensors. Applicability of FinFETs and pseudo-Hall transistors is also assessed.

Following the historical and literature overview of 3D IC technologies and CPI within 3D IC stacks, the contributions of this thesis are summarized in section 2.3.

Chapter 3

Piezoresistance effect and Stress sensor evaluation

This chapter begins by introducing the piezoresistance effect, instrumental to linking stress components and front-end-of-line (FEOL) device current shifts within this thesis. Section 3.1.1 lays out the fundamentals of the piezoresistance effect, section 3.1.2 discusses the justification of utilizing the piezoresistance effect within this thesis and section 3.1.3 discusses the versatile application of the piezoresistance model. Stress sensors used in this thesis are based on the piezoresistance effect. Following the piezoresistance effect, evaluation of stress sensors is presented in section 3.2. The developed and implemented stress sensor evaluation methodology is presented in section 3.2.1 and an introduction to the specific part of the evaluation, stress sensor calibration, is given. Stress sensor calibration continues in detail in chapter 4.

3.1 Piezoresistance effect

3.1.1 Fundamentals of the piezoresistance effect

The piezoresistance effect has been utilized in microelectronics from its discovery by Smith in 1954 at Bell laboratories [93]. It stands as one of the milestones of IC technology as through the development of piezoresistive sensors, it propelled advanced Si micromachining and paved the road to the creation of MEMS technology in the 1980s. Due to its importance and usage, many papers on its theoretical background have been written throughout the decades [94-99] and are still published today [100-104]. A timeline is presented in figure 3.1 [100] illustrating advances in IC fabrication from the 1950s to the 21st century. The application of piezoresistance emerged together with the the concept of ICs and developed hand in hand with Si advancements complementing other Si processing technologies. Today, Si piezoresistance is widely used for various sensors including pressure sensors, accelerometers, cantilever force sensors, inertial sensors and strain gauges. The usage of piezoresistance in this thesis deviates slightly from the established implementation within the MEMS industry. While keeping the concept of utlizing piezoresistance for sensorics, it applies piezoresistance to sub 100 nm and advanced technology nodes.

Piezoresistance is in essence an effect describing change of resistance of a material due to applied stress. In comparison to piezoelectricity, charges are not generated in the process of piezoresistance. In general, the electrical resistance of a homogenous structure corresponds to *eq. 3.1*

$$R = \rho \frac{l}{S} \tag{3.1}$$

where l is the lenght of the resistor, S its cross-sectional area and ρ its resistivity. Therefore, a change of resistance in a material can be a result of:

- Geometrical changes, related to the change of length l or cross-sectional area S
- Resistivity changes, related to shift of ρ value

Following that division, *eq. 3.1* can be rewritten as

$$\frac{\Delta R}{R} = (1+2\upsilon)\varepsilon + \frac{\Delta \rho}{\rho} \tag{3.2}$$

where υ represents Poisson's ratio and ε applied strain. For silicon under applied strain, the relative resistivity shift ,$\Delta\rho/\rho$, is 50-100 times higher than the geometrical term defined by l and S. Therefore, piezoresistance of silicon is based on its resistivity shift. Resistivity for isotropic materials can further be written as

$$\rho = \frac{1}{q(n\mu_n + p\mu_p)} \tag{3.3}$$

linking resistivity to fundamental values of electron and hole mobility, μ_n and μ_p, along with electron and hole concentration, n and p, and elementary charge q. From a physical viewpoint, application of stress/strain causes energy splitting in the conduction band

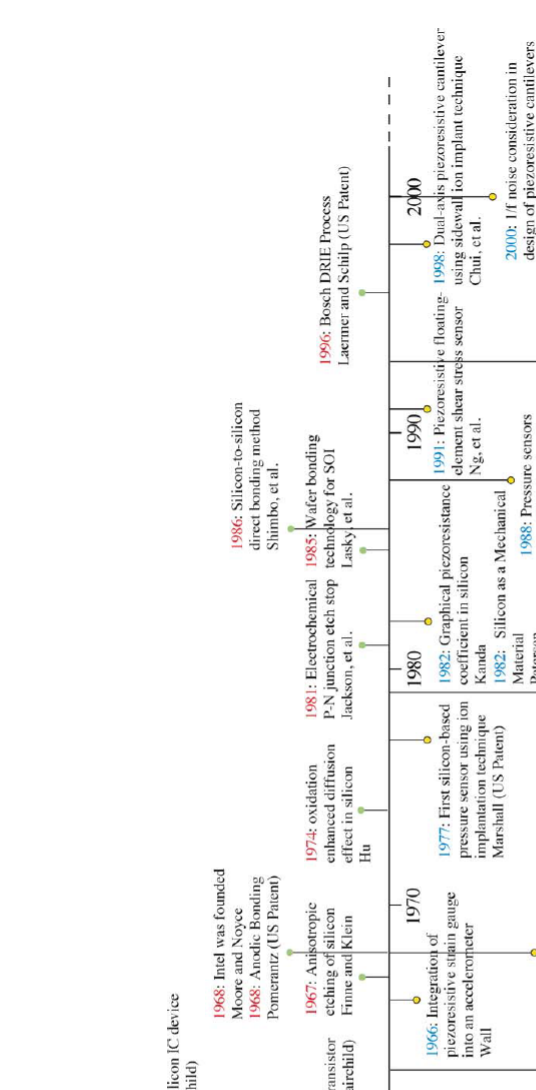


Figure 3.1: Advances in IC fabrication with piezoresistance milestones [100]

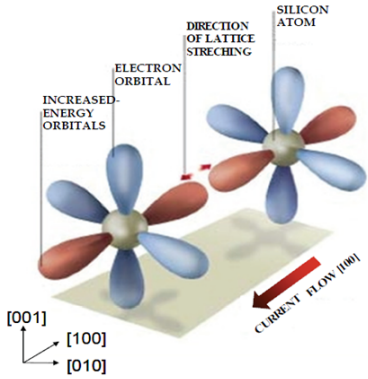


Figure 3.2: Redistribution of carriers under uniaxial stress

and shape distortion in the valence band. As indicated on an example in figure 3.2 uniaxial stress causes stretching of particular silicon orbitals. This causes a redistribution of current carriers, electrons and holes, moving to surrounding orbitals.

The mobility of carriers is described by

$$\mu = \frac{e\tau}{m^*} \tag{3.4}$$

where τ stands for carrier scattering time to another energy state, m^* its effective mass and e the elementary charge. The shift of carriers to other orbitals affects their scattering time which consequently reflects in a change of mobility. Finally, it can be said that piezoresistance in silicon is a mobility shift of carriers, electrons and holes caused by external mechanical stress. Furthermore, the mobility shift and resistivity shift can be translated to electrical current. Ideally, in bulk material the piezoresistance effect describes the impact of stress on material by a mobility shift which is directly linked to a resistivity shift and ultimately a current shift:

$$-\frac{\Delta I}{I} = \frac{\Delta R}{R} = \frac{\Delta \rho}{\rho} = -\frac{\Delta \mu}{\mu} \tag{3.5}$$

In reality, materials are not homogenous and exhibit a certain degree of anisotropy. In 1954, to mathematically model the piezoresistance effect for anisotropic materials, Smith [93] used Bridgman's tensor notation to link stress components and resistivity components via piezocoefficients. Kanda [96-98] later generalized these relations resulting in

$$\frac{\Delta \rho_\omega}{\rho} = \sum_{\lambda=1}^{6} \pi_{\omega\lambda}\sigma_\lambda \tag{3.6}$$

where π are the newly introduced piezocoefficients, σ mechanical stress with and ω and λ the subscripts Kanda chose to indicate current orientation as a function of stress. The silicon crystal lattice exhibits cubic crystal symmetry, as illustrated in figure 3.3.

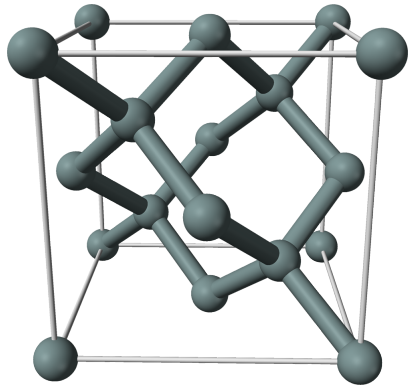


Figure 3.3: Cubic symmetry Si crystal lattice

This enables simplification of *eq. 3.6* which leads to the standard model for the piezoresistance of silicon:

$$\frac{1}{\rho}\begin{bmatrix}\Delta\rho_{11}\\ \Delta\rho_{22}\\ \Delta\rho_{33}\\ \Delta\rho_{23}\\ \Delta\rho_{31}\\ \Delta\rho_{12}\end{bmatrix} = \begin{bmatrix}\Pi_{11} & \Pi_{12} & \Pi_{12} & 0 & 0 & 0\\ \Pi_{12} & \Pi_{11} & \Pi_{12} & 0 & 0 & 0\\ \Pi_{12} & \Pi_{12} & \Pi_{11} & 0 & 0 & 0\\ 0 & 0 & 0 & \Pi_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & \Pi_{44} & 0\\ 0 & 0 & 0 & 0 & 0 & \Pi_{44}\end{bmatrix}\begin{bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\\ \sigma_{23}\\ \sigma_{31}\\ \sigma_{12}\end{bmatrix} \tag{3.7}$$

Eq. 3.7 uses the standard notation of resistivity shift $\Delta\rho/\rho$, piezocoefficients π and stress components σ in the [100]/[010]/[001] Si crystal reference frame. The Si surface is defined by [001] direction while the [100] and [010] represent in-plane Si crystal directions. The 6 independent stress components are linked to 6 independent resistivity shifts through a matrix of 3 independent piezocoefficients. Figure 3.4 visualizes 9 defined stress components and highlights the 6 independent stress components acting on corresponding planes while in figure 3.5 the stress components are placed in a setting of a Si die with a [100]/[010]/[001] reference frame. The 3 remaining stress components are equal due to shear stress reciprocity. Stresses σ_{11} and σ_{22} are normal in-plane stresses acting in the plane of the Si die surface. Stress σ_{33} is the normal out-of-plane stress acting perpendicular to the Si die surface. Stresses σ_{12}, σ_{23} and σ_{31} are shear stresses acting parallel to an observed surface. Shear stresses cause material shape deformation while normal stresses preserve the shape and change the material volume. Figure 3.5 includes an illustration of acting in-plane shear stress on the top

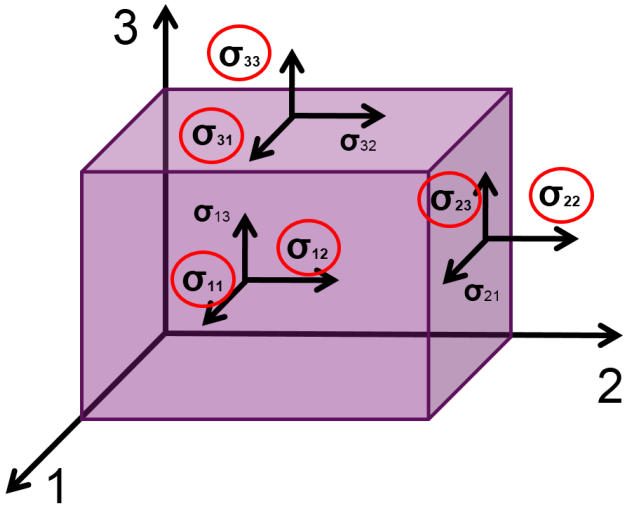


Figure 3.4: The total 9 defined stress components and the highlighted 6 independent stress components, acting on corresponding planes

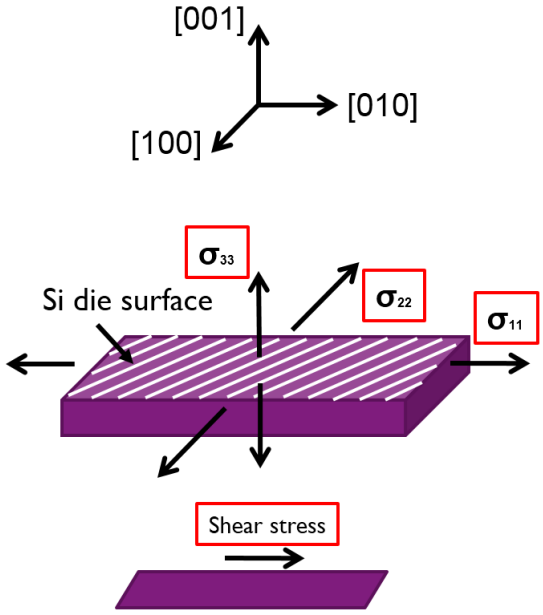


Figure 3.5: Stress components are placed in a setting of a Si die with a [100]/[010]/[001] reference frame

surface of a structure. The structure shape is deformed while the volume is preserved. Normal stresses cause deformation in a form of extension or contraction of a structure, but not its basic shape. Tensile stresses are noted with a positive sign and compressive stresses with a negative sign.

In a reference frame rotated from the [100]/[010]/[001] reference frame, the 6 independent stress components are noted as primed. Stress components σ_{11}, σ_{22}, σ_{33}, σ_{12}, σ_{23}, σ_{31} become $\sigma_{11}^{'}$, $\sigma_{22}^{'}$, $\sigma_{33}^{'}$, $\sigma_{12}^{'}$, $\sigma_{23}^{'}$, $\sigma_{31}^{'}$. Throught the thesis the stress components are attributed an alternative notation when refered to stress sensors with subscripts l, t and v corresponding to longitudinal, transverse and vertical. These subscripts were chosen to describe the orientation of the FEOL devices and applied stress relative to stress sensor current flow regardless of Si crystal orientation. This notation will be further discussed in section 3.2.2.

Piezocoefficients are the core of the piezocoefficient model and their values descibe the effect of external impact of stress on the piezoresistance of anisotropic materials. A set of piezocoefficients describes a material in any arbitrary direction. The sensitivity of the material to stress changes depending on the direction of applied stress relative to the Si crystal. Therefore, piezocoefficient values are direction dependent. Their value is further impacted by:

- temperature
- concentration of carriers

Figure 3.6 displays the varying piezocoefficients at room temperature in a (001) Si plane for n-type and p-type silicon depending on orientation of applied stress while figure 3.7 shows the impact of temperature and doping concentration to piezocoefficients in n-type and p-type silicon according to Kanda [96]. In figure 3.6, π_{long} stands for the Si piezocoefficient where stress was applied parallel to the monitored resistance shift and π_{tran} the Si piezocoefficient where resistance shift transverse to the applied stress was monitored. In figure 3.7, P(N,T) stands for the normalized value of piezocoefficients. The real piezoresistive coefficient value is calculated by multiplying the piezoresistive factor P(N,T) by the room temperature piezoresistive coefficient.

3.1.2 Justification of usage

The piezoresistance model, eq. 3.7, is an analytical linear model where stress components are related to mobility shifts and subsequently device current shifts through a system of linear equations. The model is used to characterize the electrical response of Si based devices under mechanical stress and without deeper physical insight quantifies the behavior of the stressed device. As such, it includes all possible effects to alterations of electrical current as an effect of stress, which can in nature be both linear and non-linear. Since the model is inherently linear, there are clearly limitations to its usage. A primary limitation related to non-linearity lies in the amount of applied stress. Under higher loads, the stress to mobility shift curve may harden and cease to be representable linearly. For Si stresses above 1 GPa non-linear material response could be expected.

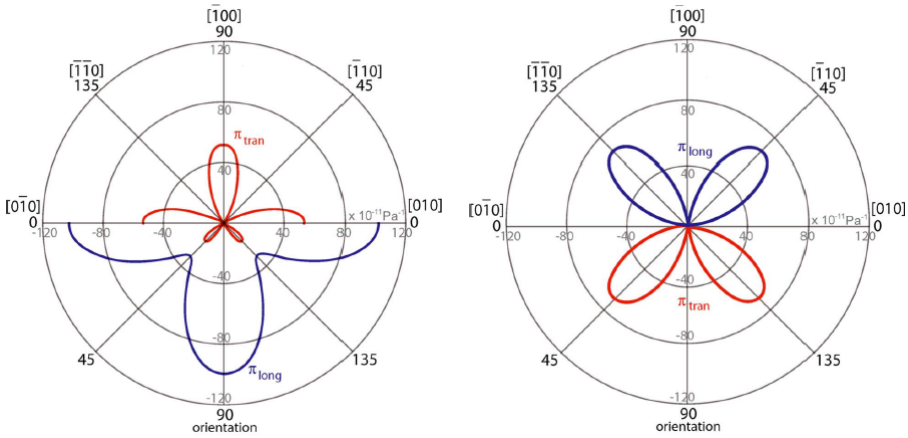


Figure 3.6: Piezocoefficients dependent on crystal orientation for a) n-type Si and b) p-type Si [96,100]

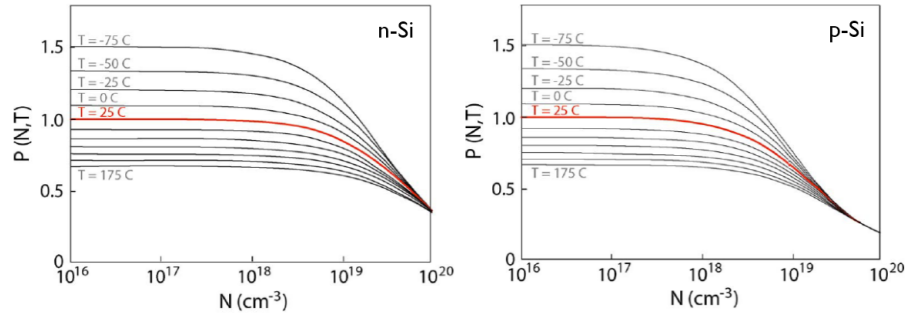


Figure 3.7: Impact of temperature and doping concentration on piezocoefficients in a) n-type Si and b) p-type Si [96,100]

As an alternative to the piezoresistance model, the S-band model can be considered. The S-band model represents a physical lower level model implementing the deformation of energy bands of a material and its immediate impact to material electrical properties. Such a model provides an in-depth analysis based on physical principles rather than the sole mathematical basis of the piezoresistance effect, but also stands as a complex representation of material behavior that can be considered a topic on its own. The piezoresistance model provides a straightforward and practical approach while maintaining desired scientific meaning.

An experiment comparing the piezoresistance model and S-band model was performed in order to gain insight in the piezoresistance model limitations. S-band simulations obtained by Synopsis Sentaurus were compared to in-house stress calibration tests. Figure 3.8 shows mobility gain responses of Si in [110] direction with current flow in the same direction and stress applied parallel, transverse and vertical to the current flow, for n-type Si, figure 3.8 a), and p-type Si, figure 3.8 b). The dashed lines represent the linear piezoresistance model while full lines represent the simulated values using the S-band model. The S-band model for both n-type and p-type Si shows non-linear behavior mostly above 1 GPa of stress. The n-type and p-type S-band curves closely follow the piezoresistance values under 1 GPa. The p-type S-band model shows more deviation from the piezoresistance model in certain cases, however the simulation was performed with a smaller amount of data points as well as exhibits a certain instability.

It can be argued that other non-linear effects may have a significant role on the final sensitivity to stress, most prominently mentioned in literature the threshold voltage of a transistor. The origin of the problem lies in the question whether or not mechanical stress has an impact on not just the mobility shifts in the Si bulk, but also on the threshold voltage of the device. If the threshold voltage would be sensitive to stress, the overall current shift of the device would not only be a consequence of the mobility shifts in the transistor channel but introduce non-linear behavior which is fundamentally not covered via the piezoresistance model. Tests regarding stress impact on threshold voltage have been performed and are presented in Chapter 4. They showed no immediate threshold voltage shift upon appliance of stress which gives further assurance for the validity of the piezoresistance model.

3.1.3 Twofoldness

The piezoresistance model has a twofold role within this thesis. It was exploited as a connector from stress to mobility shift and vice versa, mobility shift to stress. Ideally, a set of piezocoefficients can provide information from two different perspectives:

- Extraction of mechanical stress - Devices sensitive to mechanical stress can be used as stress sensors in which case their electrical responses are converted back to stress in order to obtain information on the stress fields in the surrounding and characterize a certain process
- Impact on FEOL devices - If stress is mentioned in the same sentence with FEOL, usually this connects to deliberate implementation of stress/strain within devices in order to boost their performance. This implies that the stress levels are

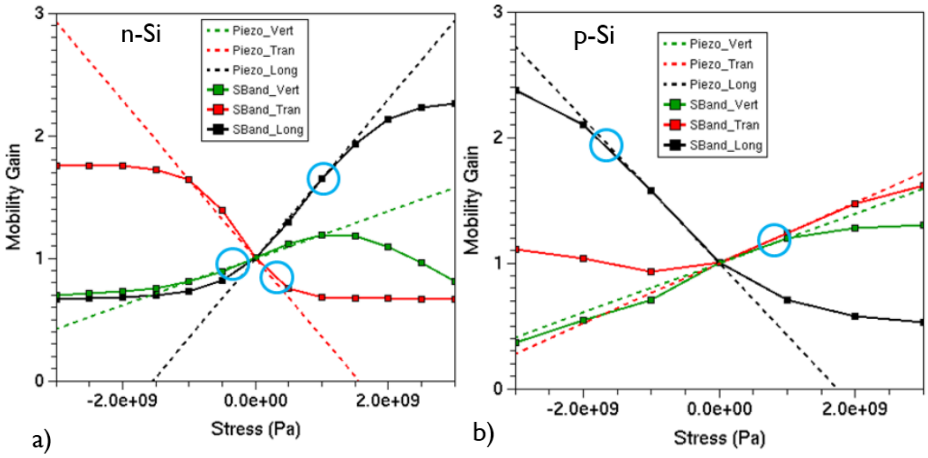


Figure 3.8: Comparison of the piezoresistance model obtained through measurements and S-band model simulated with Synopsis Sentaurus

known, the implementation process controlable and the device performance after its implementation known. Additional processing after FEOL can induce undesired additional stresses. In particular, packaging assembly phases, especially in 3D technology can have a significant impact. In that sense revealing the impact of stress on FEOL devices and their sensitivity provides valuable information on a technology node and device type. Obtaining FEOL device sensitivity to stress is of high importance as it can have significant impact on the IC layout design through introduction of proper keep-out-zones.

Chapter 4 elaborates on the findings of stress responses of FEOL devices from both perspectives, while chapters 5 and 6 include further information on the specific ways of usage of the piezoresistance model for collection of data on 3D stacking and packaging.

3.2 Stress sensor evaluation

3.2.1 Methodology

Whether a device is specifically designed to be a stress sensor or if there is potential for it to be exploited as one, a certain evaluation methodology is needed that will standardize the selection process. This methodology has to be general enough to cover a variety of sensors, their types, orientations, operation principles etc. Therefore, all devices used within this thesis that were considered as stress sensors or had potential to be used as stress sensors, underwent the same evaluation procedure until sufficient data was gathered to conclude on its usability.

The proposed and used evaluation methodology for mechanical stress sensors is schematically presented in figure 3.9 . It can be grouped in three main parts each of which consists of several tasks:

- Wafer level evaluation
  - Test vehicles
  - Stability measurements
- Stress calibration
  - Idenfitication and sample preparation
  - Calibration test
- Sensor usage
  - Validation test
  - Application

Test vehicles This methodology starts off by already implying existence of processed devices awaiting evaluation. In the early stages of the thesis existing designs already processed on particular test vehicles were taken into consideration as the idea was to attempt to determine the usability of a large group of various different devices in order to gain early knowledge on critical parameters for choice of stress sensors. Implementing new designs implied significant allocation of resources and dedication of chip floorplan solely to CPI. As at the beginning of the thesis 3D integration CPI was in its beginnings, a dedicated stress test chip was not planned until a concrete insight was established and until CPI stress sensor learning matured. Firstly, other key CPI points such as fundamental thermo-mechanical simulations and investigation of stress impact on the FEOL needed to be addressed. Furthermore, it was believed that through several learning cycles and optimization, the current list of device designs and types had potential to pave the way for CPI stress sensor methodology standardization.

The learning cycles, which were always in parallel with concrete feedback for the development of the 3D CPI program and this thesis, were performed on 3 logic

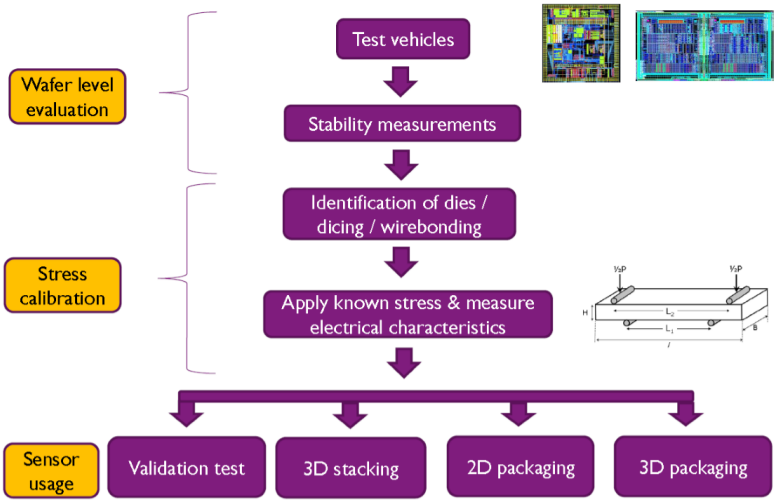


Figure 3.9: Stress sensor evaluation methodology

based test vehicles, ETNA, 3D 65 and FUJI, each a representative of its technology node, 130nm, 65nm and 32nm, respectively. Subsequently, a stress dedicated test chip, PTCQ, in 65nm technology, was designed from scratch where all the previous know-how was implemented. The test chips are discussed more in section 4.7.

The selection of devices for evaluation on these intial three test vehicles was guided by the following set of questions:

- Which stress component is addressed by these potential stress sensors (in-plane, out-of-plane, shear)?
- Where are they positioned (FEOL, BEOL)?
- Which mechanisms or assembly steps could they potentially characterize?
- What is their size?
- What is their sensitivity?

Not all questions can be upfront straightforwardly answered with certainty, especially sensitivity which is in fact the final outcome of the calibration stage, however literature on piezoresistance of Si and 2D CPI provided sound indications for commencement of device selection and evaluation.

This step also acts as preparation for electrical tests on the desired device under evaluation. This primarily includes:

- Determining device operation
- Determining device connectivity
- Defining device biasing conditions

Initial stability measurements Several factors can influence the performance of a microelectronic device, only one of which might be mechanical stress. These factors can be internal, related to the flaws in design or external, coming from the environment. Since the device under question is at this phase tested on wafer level in laboratory conditions, external factors are assumed to be negligible. These include parasitics of the wiring, contact force, electro-magnetic interference etc. All wafers were measured on a probe station with a 2x12 pin probe card as all devices were grouped in modules with 24 pads for external connection. Therefore, the fluctuations in device performance are in this case attributed to two main internal factors:

- Design quality - a device can be as good as its design allows it to be. This encompasses several steps in the device fabrication steps: fundamental knowledge, material selection, layout, processing recipes and the quality of the processing itself.
- Design tolerance - electrical current variabilities from device to device whether they are in the same module, same die, another wafer or batch are unavoidable. Invidividual design aside, each processing step operates within certain margins of precision and contributes to operational inconsistency of devices if not taken into account.

The standard procedure performed involved repeatibility measurements under biasing conditions in which values were gathered from sensors. Electrical current of each device was measured between 5 and 10 times. Regarding stress, this phase was considered as the equivalent zero stress state. If the device exhibited instability in a zero stress state, it was considered unusable and was removed from further evaluation. The exact criteria for stability evaluation depended on device type and its application. In practice, if not visibly unstable or with major flaws, the device zero stress state stability was compared to its stress sensitivity and subsequent conclusions were made.

Identification and sample preparation There is a series of steps throughout the evaluation process followed by further steps in preparation for sensor application. The number of devices can reach several hundreds per single module within one die alone. It is crucial that all tests and measurements can be correlated to their particular wafer, die, module and finally device. This necessitates software labeling of all devices during corresponding measurements as well as physically labeling prepared samples in order to ensure backtracking.

Regardless of the calibration type and tool used, isolation of individual dies by means of dicing is present. Dicing results in the transition from one whole wafer representing the distribution of samples to several tenths of standalone samples awaiting further steps for calibration. During dicing a wafer or piece of wafer is placed on sticky dicing tape held together by a metal surrounding frame. In order to decrease the chance of sample mix up after dicing, it was requested that the cut wafer was left on the dicing tape and returned as such so samples can be individually marked and prepared for further preparation requests to other parties. This proved to be a necessity in order to ensure that only the known good samples are used. Every accidently picked non-functional sample that might have found its way to the actual calibration

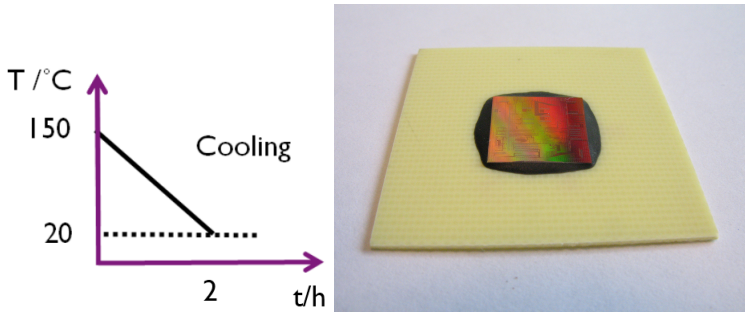


Figure 3.10: Die-to-substrate structure symbolizing the stress sensor in-plane stress validation test

implies one group of measurements lost. Ensuring high yield of samples calibrated compared to starting samples can be challenging from the technical viewpoint alone and in general the sample preparation time consuming. Therefore, it is justifiable to put in the additional effort to attempt to minimize personal and third party errors by checking the identification of the samples between sample preparation steps and in certain cases manually labeling them, even though at first glance this seems to be redundant. A misunderstanding in sample identification can potentially lead to repetition of the whole evaluation cycle. The particular sample preparation for calibrations will follow in further subchapters, 3.3.2 and 3.3.3, respectively, together with the description of the corresponding tool while subchapter 3.2.2 deals with calibration notation. The tools used for stress sensor calibration to specific stress components are:

- 4-point bending - in-plane and out-of-plane stress calibration
- nanoindentation - out-of-plane stress calibration

Sensor usage When a stress sensor was deemed viable for implementation, on one hand, it was automatically considered for inclusion in the future dedicated stress test chip, PTCQ, and on the other, the current test chip was forwarded for submission to 3D stacking or 2D packaging. The ETNA test chip was submitted to 2D packaging and 3D stacking, the FUJI test chip to 3D stacking while the PTCQ test chip was submitted to all 3D assembly phases, 3D stacking and 3D packaging, along with 2D packaging for gaining an overall 2D vs 3D picture.

Prior to 3D stacking or packaging application, a final test was performed referred to as the validation test. The aim of this test was to simulate a realistic stress environment for the stress sensors but using a simple enough environment which can be relatively easily confirmed by other techniques. The comparison of the results obtained from the stress sensors with results given by other techniques benchmarked the stress sensor operation and provided a final validation of its usability.

The validation test was fully developed for in-plane stress sensors using a single chip sattached to a substrate, as in figure 3.10. Functional dies were selected through initial wafer level electrical measurements and submitted to dicing. Each die was

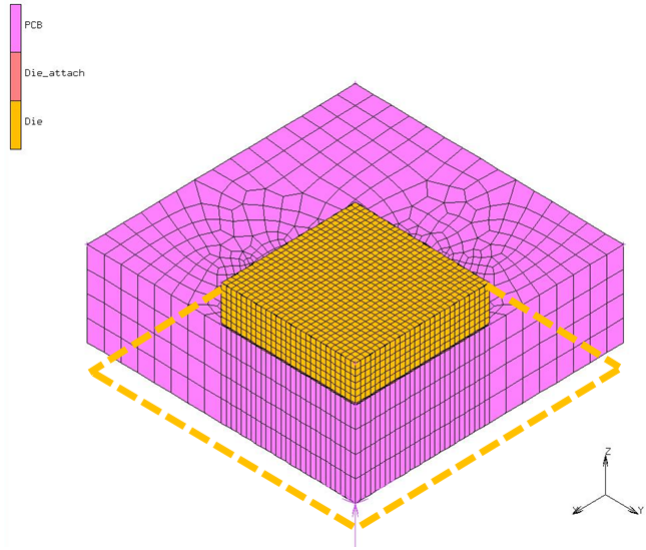


Figure 3.11: Quarter symmetry finite element model of the die-to-substrate structure

square shaped and attached at room temperature to a square shaped substrate twice the size of the die by means of a die attach, in such a way that the sensors on the die were facing upwards and positioned in the center of the substrate, die edges parallel to substrate edges. Such a still loose structure was placed in the oven for die attach hardening resulting in bond stiffening. During cooling of the structure, the higher CTE of the substrate, compared to Si CTE, caused convex bending of the die-to-substrate structure which resulted in in-plane stresses on the surface of the die. The electrical results gathered from the sensors, before and after this die bonding, were compared to results obtained by finite element modeling and interferometric profilometry measuring the die warpage. The final die-to-substrate structure is shown in figure 3.10, where a 10.5x10.5 mm full thickness 700 μm ETNA die is glued to a 0.8 mm plastic substrate. The selected die attached necessitated 2h of baking at 150ºC followed by cooling to room temperature.

The development of the die-to-substrate structure involved two main action points:

- Choosing die substrate material and defining its thickness
- Assessing the impact of die attach thickness on stress levels within the Si die

The substrate material and thickness was chosen based on parametric finite element models. The main selection criterion was production of detectable stress levels on the Silicon surface, preferably above 100 MPa. The proposed substrate materials were plastic and copper. The finite element model of the die-to-substrate structure with only a quarter of the structure simulated because of inherent symmetries, is presented in figure 3.11, while the material properties used are shown in table 3.1. Values are taken from datasheets apart from the plastic substrate which was approximated. All materials

FEM material properties	Young's modulus [MPa]	Poisson ration	CTE [ppm/K]
Die	169000	0.3	2.6
Die attach	4900	0.3	30
PCB	20000	0.3	17
Cu	117000	0.3	17.6

Table 3.1: Die-to-substrate finite element model material properties used

Si stress [MPa]	0.1 mm	0.2 mm	0.8 mm	1.6 mm	2.4 mm	3 mm	6 mm	12 mm	24 mm
Plastic			+106					-114	-147
Cu	+98	+137	+113		-100	-142	-244	-292	-321

Table 3.2: Parametric study of substrate material and thickness selection. Maximum in-plane stress from the center of the top Si surface versus substrate material and thickness.

were modeled as isotropic. Table 3.2 presents the obtained stress results in Si, extracted from the center of the top side Si surface. A plastic substrate of 0.8 mm thickness was chosen as it provided representable stress levels and was easier to section in pieces from bulk compared to copper. Figure 3.12 contains top side Si surface stress patterns represented in quarter symmetry. The two in-plane stresses are distributed uniformly across the whole die and change only close to its edges while the shear stress and out-of-plane stress is insignificant.

3.2.2 Calibration

Calibration of stress sensors presents one of the fundamental aspects of this thesis. This is in a large majority done with the 4-point bending setup. The outcome of the calibration are the stress sensitivity coefficients referred to as piezocoefficients, explained in the previous chapter of piezoresistance. Obtaining a particular piezocoefficient provides the link between a certain stress component and the current shift of the FEOL device. As initially mentioned in chapter 3.1 it is assumed that under moderate levels of stress in silicon (approximately under 1 GPa) the current shifts are proportional to a linear combination of acting stress components, therefore eligible for representation via the linear piezoresistance model, eq. 3.7. The piezoresistance itself describes the linear dependency of the relative resistivity shift of a material to mechanical stress. The relative resistivity shift basically originates from the relative mobility shifts of electrons or holes, respectively. In pure silicon, the relative mobility shift can be directly related to current shift, however in FEOL devices, the current shift can potentially be impacted by several other factors, threshold voltage shift being the most mentioned one. In general,

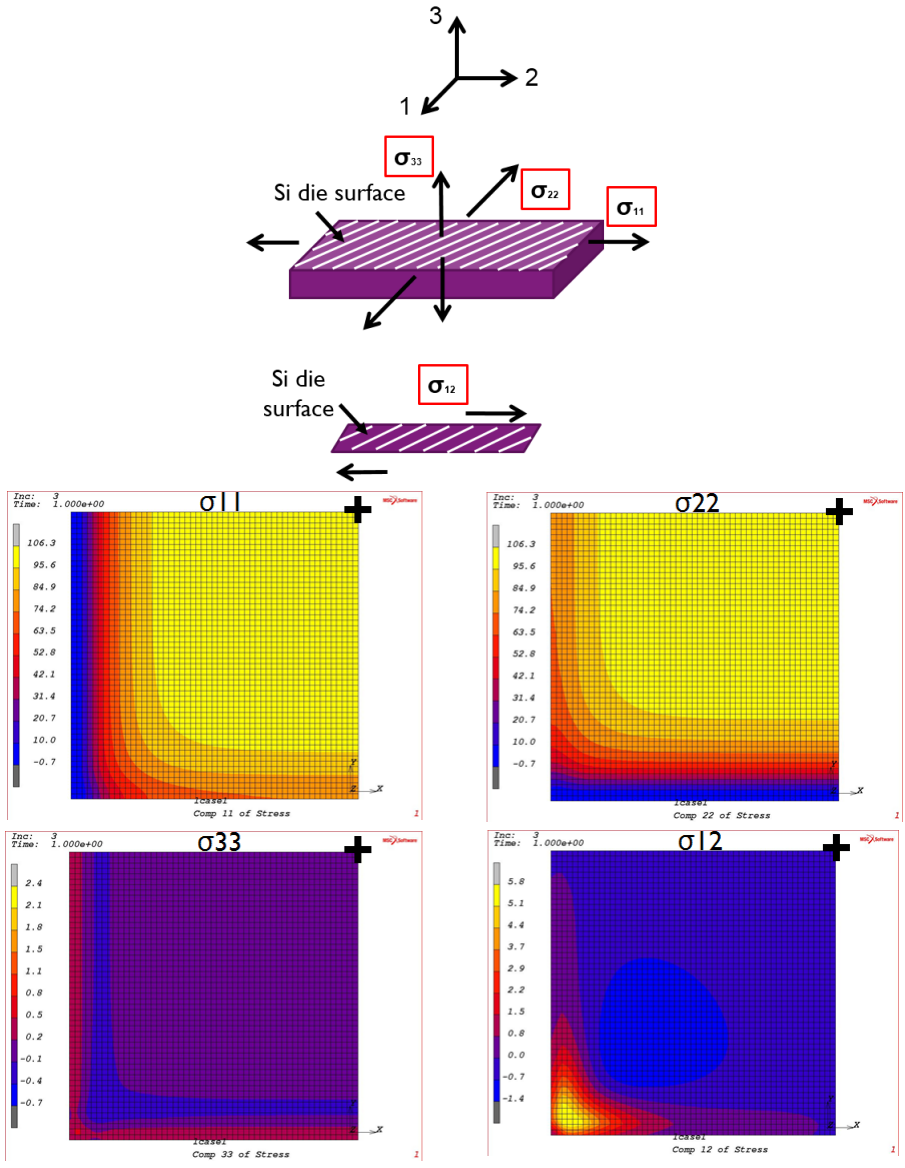


Figure 3.12: Quarter symmetry representations of in-plane, out-of-plane and shear stress on the top Si surface

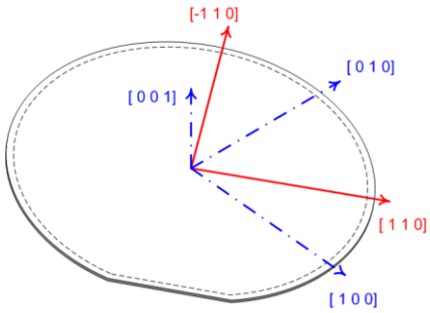


Figure 3.13: Si crystal orientations on a Si wafer. A Si wafer typically has a [001] surface with devices oriented 45º from the standard Si crystal orientations. This places the FEOL devices with current flow in either [110] or [-110] direction.

when substituting relative mobility shift with relative current shift in piezoresistance formulas, piezocoefficients become stress sensitivity factors which include all factors with linear impact to current shift caused by mechanical stress which are not known at the moment of calibration. Therefore, even without an in depth analysis of the possible contributions to current shift, besides mobility shifts and threshold voltage shifts, the piezocoefficients are a reflection of the real stress state the FEOL device finds itself in and the particular electrical response to surrounding stress.

If we consider the typical Si wafer surface to hold [001] orientation and the two most typical orientations for FEOL devices to be [110] and [-110], as indicated on the schematic in figure 3.13, the piezoresistance effect for devices in these two orientations notated with respect to the crystal directions of importance can be described with the following sets of functions:

$$\frac{\Delta I_{110}}{I_{110}} = \pi_{110}\sigma_{110} + \pi_{-110}\sigma_{-110} + \pi_{001}\sigma_{001} \tag{3.8}$$

$$\frac{\Delta I_{-110}}{I_{-110}} = \pi_{-110}\sigma_{110} + \pi_{110}\sigma_{-110} + \pi_{001}\sigma_{001} \tag{3.9}$$

$\Delta I_{110}/I_{110}$and $\Delta I_{\text{-}110}/I_{\text{-}110}$represent the relative current shift responses for devices with current flowing along the [110] and [-110] axes, respectively, as a consequence of the in-plane stresses $\sigma_{110}$and $\sigma_{\text{-}110}$ acting along the same axes and out-of-plane stress σ_{001}acting along the [001] axis, perpendicular to the Si wafer surface. The type and severity of impact is determined by the corresponding linking piezocoefficients π_{110}, $\pi_{\text{-}110}$ and π_{001} .

With respect to the impact of stress on FEOL devices, a more convenient notation than the one from *eq. 3.8* and *eq. 3.9* is used throughout the thesis. Subscripts *l*, *t* and *v* standing for longitudinal, transverse and vertical are used, respectively. This notation refers to the orientation of the FEOL device and direction of stress relative to the FEOL device current flow. An orientation of a device in, for instance, [100] or [110] direction implies its current is flowing in [100] or [110] direction, respectively. If the initial direction of the FEOL device current flow is labeled as longitudinal, a

direction 90º from this one in the plane of the Si surface can be labeled transverse and a direction 90º perpendicular to the plane of the Si surface can be labeled vertical. For a FEOL device with current flow in longitudinal direction, stress σ_l is then the stress acting parallel to the device current, stress σ_t is the stress acting transverse to the device current in the plane of the Si surface and stress σ_v is the stress acting perpendicular to the device current normal to the plane of the Si surface. Piezocoefficient π_l then links the impact of stress σ_l to current shift, piezocoefficient π_t then links the impact of stress σ_t to current shift and piezocoefficient π_v then links the impact of stress σ_v to current shift. These notations are meant to be used regardless of the Si crystal reference frame.

For instance, in case of a [100]/[010]/[001] reference frame and a device current flow in [100] direction, [100] would be the longitudinal direction, [010] the transverse direction and [001] the vertical direction. In case of a 45º rotated reference frame, [110]/[-110]/[001], with device current flow in [110] direction [110] would be the longitudinal direction, [-110] the transverse direction and [001] the vertical direction. Therefore, *eq. 3.8* and *eq. 3.9* in its adapted form in which they are reffered to throughout the thesis look like:

$$\frac{\Delta I_1}{I_1} = \pi_l \sigma_l + \pi_t \sigma_t + \pi_v \sigma_v \tag{3.10}$$

$$\frac{\Delta I_2}{I_2} = \pi_t \sigma_l + \pi_l \sigma_t + \pi_v \sigma_v \tag{3.11}$$

The relative current shifts are chosen to be notated with digits 1 and 2 to keep in mind at all times that two separate devices are in question.

3.3 Summary

The piezoresistance effect is a fundamental effect in Si linking mechanical stress to resistivity and resistance shift in Si. In essence, applied mechanical load impacts the mobility of current carriers, electrons and holes. A linear model considers anisotropy of resistivity in Si and connects 6 independent stress components, in-plane stress, out-of-plane stress and shear stress, to material resistivity shifts over a matrix of piezocoefficients. Piezocoefficients define the stress sensitivity of Si in arbitrary directions. Piezocoefficient values change with direction of applied stress, doping levels in Si and temperature. Related to direction of applied stress, the resistivity shift is impacted by the relative direction of the applied mechanical load to the resistivity under observation.

Within this thesis, the linear piezoresistance model, originally developed for bulk Si, is applied to describe the stress behavior of FEOL devices. In this case, the piezocoefficients represent the stress sensitivity of the whole device. For Si based FEOL devices, such as MOSFETs, their drain to source current is directly proportional to electron or hole mobility. The piezoresistance model applied to FEOL devices now links independent stress components, with device piezocoefficients, to device current shift. Although with application of the piezoresistance model FEOL devices, it is still assumed that current shifts are primarily caused by mobility shifts in device channels,

other impacts to current shift can be included. The piezoresistance model can be considered valid for FEOL devices as long as all impacts as a consequence of mechanical load, contribute linearly to the final current shift.

As an alternative to the piezoresistance model, the S-band model can be considered. The S-band model represents a more physical model encompassing the deformation of energy bands of a material and its immediate impact to material electrical properties. Such a model provides an in-depth analysis based on physical principles, but also stands as a complex representation of material behavior that can be considered a topic on its own. The piezoresistance model provided a straightforward and practical approach while maintaining desired scientific meaning.

A stress sensor evaluation methodology is proposed and implemented within the thesis. Its 3 phases are wafer level evaluation, stress calibration and sensor usage. During wafer level evaluation, the stability of the device is assessed with no external loads applied. If the device exhibits sufficient electrical stability, it can proceed to the stress calibration phase. Stress calibration involves current shift monitoring during application of a controlled external mechanical load. Device piezocoefficients can then be extracted as a result of stress calibration which describe the sensitivity of the device to in-plane, out-of-plane and shear stress. If the device exhibits sufficient sensitivity to stress it can be submitted to a validation test. The validation test, where possible, is the final step before sensor application in a real environment. The Si die with processed stress sensors is placed in a simplified and known stress environment. Stress is extracted with stress sensors and compared to other stress extraction methods using techniques such as profilometry or finite element modeling.

In-plane and out-of-plane stresses are noted throughout the thesis in 2 different ways. When stresses are generally discussed or in terms of a finite element simulation, they are linked to the notation by Smith. In a standard Si crystal [100]/[010]/[001] reference frame the two in-plane stress components and out-of-plane component are refered to as σ_{11}, σ_{22} and σ_{33}, respectively. In an rotated reference frame, such as the [110]/[-110]/[001] reference frame frequently used in microelectronics rotated 45º from the standard Si crystal refrence frame, the in-plane stress components and out-of-plane stress component are refered to as $\sigma_{11}^{'}$, $\sigma_{22}^{'}$ and $\sigma_{33}^{'}$, respectively.

When stress components are refered to the stress sensors, since sensors are often placed 90º degrees to each other, in longitudinal and transverse direction, the in-plane stress components and out-of-plane stress component are refered to as σ_l, σ_t and σ_v, respectively. The corresponding piezocoefficients then follow this notation as π_l, π_t and π_v, respectively. The subscript l stands for longitudinal, t for transverse and v for vertical corresponding to the orthogonal reference frame the sensor is placed in.

Chapter 4

Methods and techniques

In this chapter, all the methods, techniques and their corresponding tools are presented. Firstly, a basic description of each utilized technique is given followed by a section on their application within this thesis. Sample preparation is summarized related to each method and technique, where applicable. The chapter discusses in order of appearance: the finite element method (FEM), 4-point bending, nano-indentation, optical and mechanical profilometry and X-ray diffraction.

4.1 Finite element method

4.1.1 Basic description

”The limitations of the human mind are such that it cannot grasp the behaviour of its complex surroundings and creations in one operation. Thus the process of subdividing all systems into their individual components or ’elements’, whose behaviour is readily understood, and then rebuilding the original system from such components to study its behaviour is a natural way in which the engineer, the scientist, or even the economist proceeds” [105]. With these sentences O.C. Zienkiewicz, one of the original authors of the finite element method, and R.L. Taylor, start their introduction to the method that has from its origins more than 40 years ago until today become fundamental in simulation softwares in many industrial branches from aerospace, automotive, manufacturing, energy to electronics, chemical, consumer and other modern industries [106].

A sequential summary of modeling the continuum with a finite element process follows:

1. Continuum discretization and node assignement - diving the continnum into finite elements and defining the connection between integration points and nodes of elements
2. Element matrixes - building matrix equations expressing the properties of indivitial elements
3. Assembly - Obtain system equations by assembling the overall system of individual element properties
4. Boundary conditions - Modifying system equations by taking into account the imposed boundary and loading conditions
5. Solution - Processing the set of simulatenous equations in order to calculate the nodal values of the problem
6. Post-processing - Often needed to calculate other parameters connected to the main equation solutions

Application of the finite element method to a problem produces a representation of the real state and acts as an approximation of the actual structure behavior. With this in mind, it has to be said that the realism of the final results depends on the quality of the built model. There are several aspects of the model that are critical to result credibility:

1. Meshing - During continuum discretization, areas with higher gradients are more critical than others and require denser meshing for more accuracy in the particular region. Furthermore, the type of elements used and their shape after meshing can change the final values which implies they need to be addressed accordingly.
2. Material properties - Even if the mesh is optimal, the degree to which the structure will react to its environment is defined through the properties of constituent materials. There are several ways of obtaining material properties for the model, from pure assumptions if material data is not known to datasheets or in-situ material measurements.

3. Loading and boundary conditions - Since the meshed structure is initially in free space, applying boundary conditions controls the degrees of freedom of the structure and with it its reaction to the applied loads. The right boundary conditions can also cut the size of the model considerably saving valuable computation time. When considering the applied loads, the final results are mere direct outputs excited by the initial conditions and applied stresses to the structure. Obtaining more fundamental details about the process which is being modeled and thereafter the choice of the load type, way of application to the structure and the value of the load need to be carefully considered.

4. Solution algorithm - The finite element model is in essence a mathematical model designed to solve a complex set of differential equations. The algorithm used to handle this large system, i.e. variations of the Newton-Rhapson procedure and additional options available to the user, i.e. number of simulation steps, can prove to be the difference in model convergence and substantially impact the model resource and time consumption.

4.1.2 Usage in this thesis

All simulations performed in this work were conducted using MSC.Marc commercial finite element method software [107]. MSC.Marc is used as the solver, while MSC.Mentat is used as the pre and post-processing software. Its primary usage in this thesis lies within the field of thermo-mechanical interaction of materials where the solutions to problems represent stress and strain distributions within a continuum.

Selected features of the finite elements models built in this thesis are mentioned in the following lines:

- Representation of results - The standard output of finite element models in this thesis are mechanical stress components, as indicated in figure 3.4. The notation of the stress components is discussed in section 3.1.1 and 3.2.2. When referring to finite element models, stress components are noted with the digit representation as in figure 3.4, while when referring to stress sensors, the normal stress components are noted with letters l, t and v, as in eq. 3.10 and eq. 3.11. Apart from individual stress components, major principal stress is also used to point out stress concentrations within a structure. The major principal stress outputs a maximum absolute value of the normal principal stresses.

- Usage of subroutines - Apart from reviewing stress related results, in order to implement the piezoresistance effect subroutines are used to calculate current shifts in N-type and P-type silicon from the surrounding stress state. The equations implemented standardly neglect the impact of out-of-plane stress unless otherwise noted. Furthermore, sets of equations are implemented for both standard device orientations, [110] and [-110], respectively. The basic set of equations for an arbitrary crystal orientation are presented below, where σ_l and σ_t respresent longitudinal and transverse in-plane stress in the direction of and perpendicular to the current flow I , together with their corresponding longitudinal and transverse piezocoefficients, π_l and π_t, respectively.

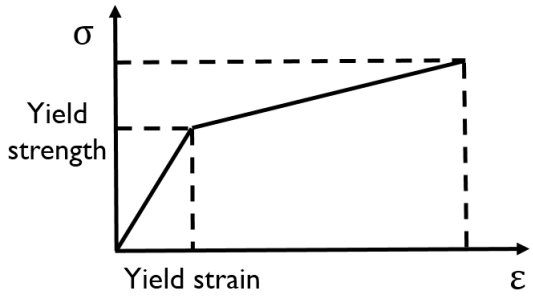


Figure 4.1: Idealized stress/strain curve implemented for copper

$$\frac{\Delta I_N}{I_N} = \pi_{Nl}\sigma_l + \pi_{Nt}\sigma_t \tag{4.1}$$

$$\frac{\Delta I_P}{I_P} = \pi_{Pl}\sigma_l + \pi_{Pt}\sigma_t \tag{4.2}$$

- Modeling material properties - Without getting into the calculation algorithms behind the software, two properties of materials used in finite element models should be mentioned, material uniformity and plasticity. Firstly, silicon was modeled as isotropic in the models where out-of-plane stress/strain was considered not to play a vital role and orthotropic where the out-of-plane component was thought to have a tangible impact to the final values. Furthermore, linear elastic models used extend to the non-linear regime in situations where plasticity is applied for copper. Plasticity is implemented with a piecewise linear function, an idealized work hardening stress/strain curve, as shown in figure 4.1.

4.2 Delaminator

4.2.1 Basic description

The delaminator is a high precision micro-mechanical test system with a primary function of providing controlable mechanical tension or compression. Simple in nature, light, portable in size and with straightforward operation, it is a practical and widespread tool particularly in material science, extending its reach further to microelectronics as well. The standard delaminator setup used is a DTS (Dauskardt Technical Services) delaminator test system with full computer control and data analysis. Figure 4.2 shows the delaminator with its main parts. The heart of the tool are the ultra high resolution linear actuator and load cell, where the actuator provides and monitors displacement towards or away from the load cell and the load cell itself stays fixed monitoring the applied mechanical load. The load cell and actuator are enforced with two steel pillars grasped around by cubic metal weights, one of which is moveable to enable an arbitrary

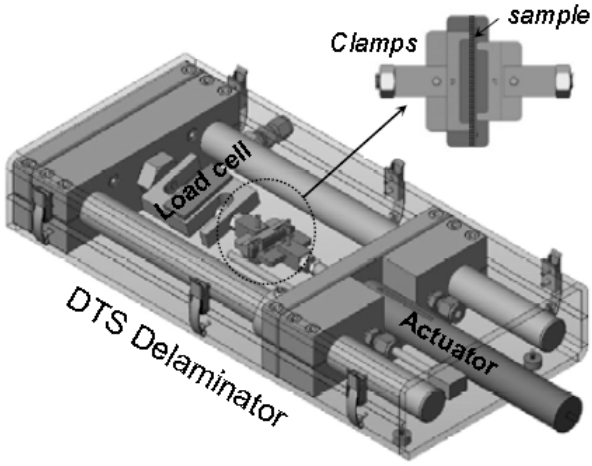


Figure 4.2: Delaminator illustration [108]

starting position of the actuator. This mechanical stiff frame plays a role in improving test stability and measurement yield.

Figure 4.3 presents the utilized delaminator setup, with the delaminator tool in the bottom right corner and according supporting equipment and computer control in the top left. The possible speed of the actuator spans from 20 nm/s to 3 mm/s, while the maximum withstandable load on the load cell is 222 N. For safety reasons, the load on the actuator is software limited to 180 N. This load can only be applied in the direction of the displacement axis. The specified operational temperature range is from -15ºC to 65ºC.

The delaminator can be used to perform a great variety of tests and apart from its straightforward operation, it owes its diversity to the exchangable clamps. The clamps are attached separately, mostly do not require extensive design time, replaced easily leading to fast adaptation for subsequent test enablement.

The most common usages of the delaminator are the following:

- Adhesion tests
  - mode I fracture: Double Cantilever Beam (DCB)
  - mode II fracture: End Notch Flexure test (ENF)
  - mixed fracture: 3-point or 4-point bending test (3-4PB)
- Ultimate tensile strength test (UTS)
- Young's modulus test

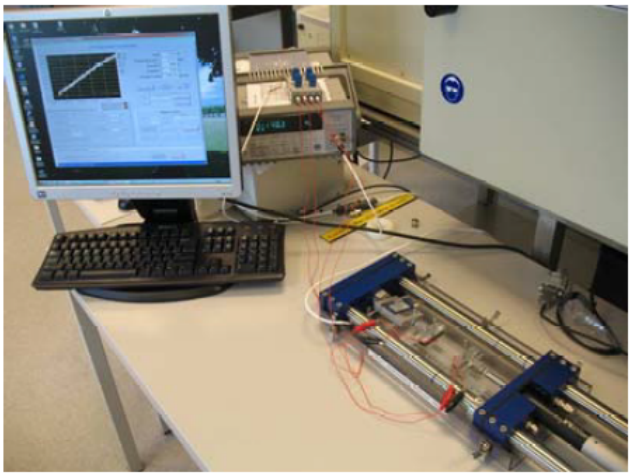


Figure 4.3: Delaminator setup and its components stationed in the lab

4.2.2 Usage in thesis

Within this thesis, the delaminator is used as an in-plane and out-of-plane stress sensor calibration tool. The following paragraph explains the sample preparation for 4-point bending and adhesion tests followed by its usage as DCB tool.

4.2.3 4-point bending

4.2.3.1 Sample preparation 4-point bending stress calibration

The delaminator is used as the 4-point bending tool for the following calibrations:

- Calibration of in-plane stress
- Calibration of out-of-plane stress

In-plane stresses imply the ones acting in the plane of the die surface, while out-of-plane implies the direction perpendicular to the die surface.

In-plane stress calibration In-plane stress calibration requires a fairly lengthy sample preparation process and follows the following main steps:

1. Die selection
2. Dicing
3. Wirebonding
4. Test setup

Once a device has been cleared for calibration after the stability measurements, particular dies and their total number have to be chosen. The total number of calibrated devices and in that sense the number of dies chosen is a trade off between statistics and total preparation and measurement time. The nature of the setup is such that it necessitates samples usually larger than the one die size. This is due to the dimension of the 4-point bending clamps where the larger of the two are on a distance of 30 mm.This means that in order to apply adequate in-plane stress samples need to be larger than 30 mm in length. The inner clamps stand apart 15 mm, which implies that devices under calibration and area covered by wirebonds to them should not exceed 15 mm in length, however this dimension was in practice never an issue. The height of the clamps is 10 mm which implies there is a risk of unadequately applied in-plane stress and local stress anomalies if the sample is considerably higher than the clamps.

Dicing is performed in such a way that the total silicon sample is larger than 30 mm in length and that die scrib lines are followed. Usually for each device under calibration horizontal and vertical samples are made. These are used for stress applied in parallel with the current flow of the device, referred to as longitudinal devices, and perpendicular to the current flow of the devices, referred to as transverse devices. An optimal dicing pattern needs to be made on the wafer or part of wafer designated for device calibration.

Once dicing is finalized, each sample consists of approximately 3 to 5 dies, with the middle die holding the modules and devices of interest. After isolating the samples from the wafer the selected module or modules need to be electrically connected to the outside world while at the same time maintaining mechanical robustness of the whole sample for later bending. These demands are met through the process of wirebonding where up to 24 wirebonds are made from the die pads to 24 pads on one side of a flexible substrate. Conductive lines run through the flexible substrate to its other end where its 24 pads are fixed into a socket. Prior to the wirebonding process, this flexible substrate needs to be glued to the non-active back side of the Si sample. The socket in which the flexible substrate is tucked into is a part of a PCB to which external SMUs are connected for biasing and sensing.

An example of a sample made, with 5 dies in a Si strip, wirebonded to the flexible substrate and placed into the socket is presented in *figure 4.4*. A magnification of the wirebonding area is shown in *figure 4.5*., where a 2x12 module is fanned out to the 24 pads of the flexible strip. Such a Si strip is placed in between the bending clamps of the 4-point bending tool, as shown in *figure 4.6*. The alignment of the sample with the clamps is done manually, such that the PCB is vertically adjusted to the level where the Si strip concides with the clamp position and horizontally adjusted so that die of interest is within the boundaries of the inner clamps. The in-plane stress calibration commences with electrical monitoring of the wirebonded devices during bending of the Si strip with help of the 4-point bending clamps. The whole process is rather tricky and if not done carefully, can result in shorts or opens of the wires.

The delaminator records the force-displacement curve from the sample bending. Applied in-plane stress is calculated as sample bending stress, according to *eq. 4.*3.

$$\sigma = \frac{3(L_2 - L_1)}{2BH^2} P \tag{4.3}$$

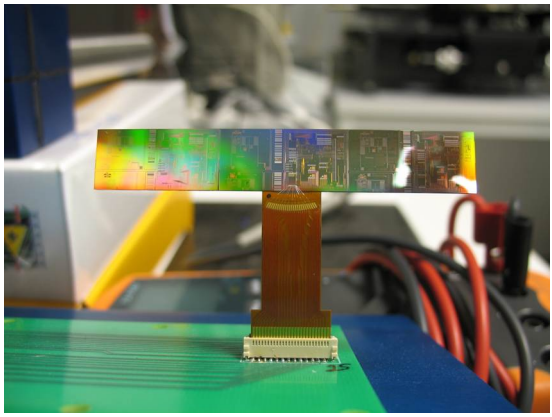


Figure 4.4: Prepared 4-point bending sample for stress calibration

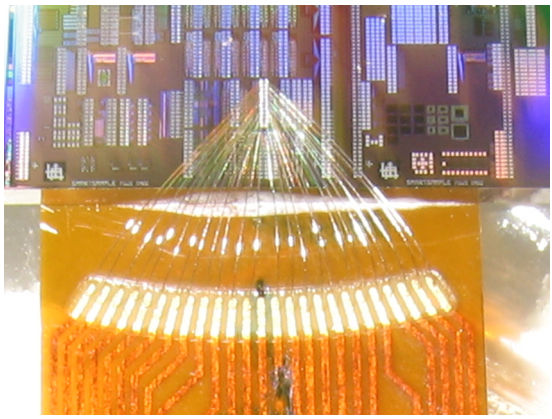


Figure 4.5: Magnified view of 4-point bending sample wirebonding area

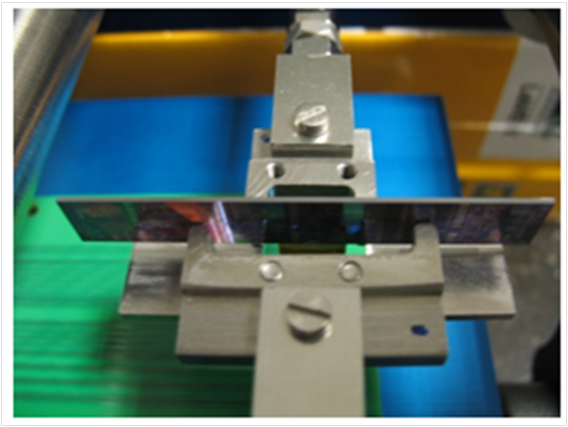


Figure 4.6: 4-point bending sample between the clamps of the delaminator

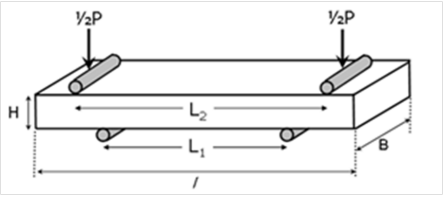


Figure 4.7: 4PB in-plane stress calibration illustration with according markage of geometry necessary for stress extraction [108]

where P represents the applied force, L2 and L1 the distances between the outer and inner clamps, respectively, H the thickness of the sample and B the width of the sample, as illustrated in figure 4.7.

Out-of-plane stress calibration Calibration of a device to out-of-plane stress requires the same steps as elaborated in the previous chapter during in-plane stress calibration, with particular differences in the dicing, wirebonding and test setup stages. Firstly, dicing is performed to isolate individual dies. These dies are placed in fan-type ceramic lidless through hole packages and wirebonded to the surrounding pads on the package lead frame. Such a package is placed within a socket to which external connections for SMUs are made. The socket with a package fixed inside is placed on a chuck type platform which represents one of the clamps of the 4PB tool. The other clamp consists of an aluminium cyllindric pillar with which force was applied to the sample. A silicon cube, dimensions 2.5x2.5x0.8 mm was placed between the Si die and the cyllindric pillar for even distribution of stress on the die area below the cube. Therefore, stress was applied by displacement of the cyllindrical clamp and application of force on the Si cube lying on the area above devices of interest for calibration. *Figure 4.8* depicts the overall 4-point bending out-of-plane calibration setup and fore-

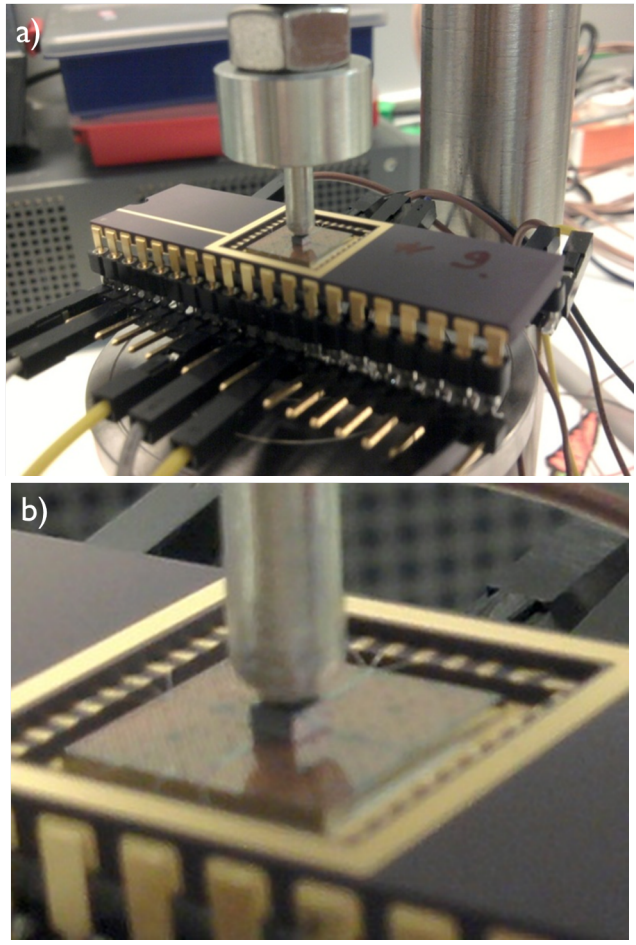


Figure 4.8: a) Out-of-plane stress sample calibration and b) magnification of the area around the metal pillar and Si cube

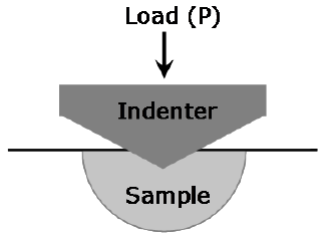


Figure 4.9: Illustration of an indenter pushing into the sample [109]

mentioned parts. For this purpose, the 4PB setup was placed vertically so that the sample lies horizontally while the cyllindric clamp applies force through downward motion perpendicular to the die.

4.3 Nano-indenter

4.3.1 Basic description

Indentation is the most commonly used technique to extract mechanical properties of materials. The advantages of carrying out indentation tests are its simplicity and speed of execution. In essence, a hard tip, of which mechanical properties are known, is pushed into a sample whose properties are unknown, as indicated in *figure 4.*9. Based on the known properties of the tip and leftover indentation mark on the sample, material mechanical properties are calculated.

A nano-indenter is a more sophisticated advancement of a regular indenter. Its purpose is to extend the ability of extraction of mechanial properties of materials in the nanometer and micrometer range and to provide information on material property changes with size. Samples are measured in small material volumes while high resolution instruments are used to control and monitor the indenter applier force and resulting displacements in the materials. A standard output of the nano indenter and a basis for calculation of the desired material properties is the load-displacement curve, presented in *figure 4.*10, obtained during both indenter intrusion and withdrawal from the material. The load displacement curve enables obtaining material properties without imaging the leftover indentation mark. The most frequently nano indenter extracted material properties are hardness and Young's modulus.

4.3.2 Usage in this thesis

The nano indenter used was an MTS Nano Indenter XP. Its main components are presented in *figure 4.*11. and listed below:

1. MTS Nano Indenter XP
2. minus k vibration isolation table

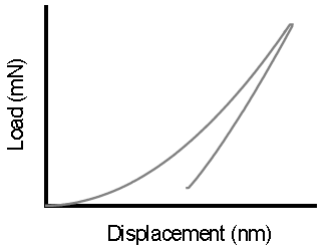


Figure 4.10: A typical load-displacement curve obtained by the nano indenter [109]

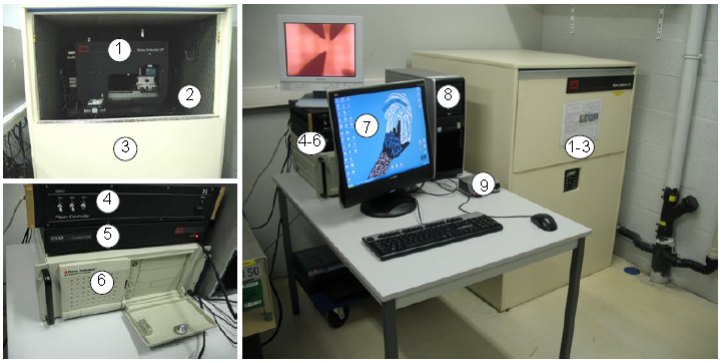


Figure 4.11: Main components of the MTS Nano Indenter XP setup

3. vibration isolation cabinet
4. Piezo controller for the NanoVision option
5. Continuous Stiffness Measurement control unit
6. data acquisition/control unit
7. monitor
8. CPU
9. Illuminator

Due to the high sensitivity of the system to the environment, parts 2 and 3 act as vibration isolators and in addition, ensure thermal and sound protection. The NanoVision option enables 3D profilometry on a micrometer range area primarily used for pinpointing the position of the indenter tip, while the Continous Stiffness Measurement control unit represents supporting hardware for an option of the tool to measure material hardness as a function of depth.

The nano indenter setup consists of two indentation systems, the standard XP indentation system and the dynamic contact module (DCM). The basic difference between these two systems is in their performance. The DCM system is more precise

MTS Nano Indenter XP specifications	Standard XP	DCM
Displacement resolution	20 pm	0.2 pm
Maxiumum indentation depth	500 μm	15 μm
Load resolution	500 nN	1 nN
Maximum load	500 mN	10 mN

Table 4.1: Main specifications for the MTS Nano Indenter XP setup

Figure 4.12: Spherical sapphire tip

in force and displacement control but with lower maximum load. The main specifications for the MTS Nano Indenter XP and its two indentation systems are presented in *table 4.1*. Due to the necessity of applying higher forces during the vertical calibration of devices than the ones managable by the DCM system and the indifference towards specified resolution, the standard XP system was chosen for measurements.

4.3.3 Sample preparation

The ease of usage of the nano indenter extends to the simplicity in sample preparation as well. A sample, small piece of wafer or die, is mounted on a metal stand of several cm across via an adhesive. Mostly a simple epoxy is heated up on the metal stand at approximately 120 ºC followed by sample placement. Alternatively, ordinary candle wax can be used for adhesion and melted on the stand with temperatures around 60 ºC. The heat transfer and mechanical effects during sample mounting are negligible. Prior to indentation, an appropriate indentation tip has to be selected. For vertical calibration tests, a spherical 250 μm in diamater sapphire tip was used, shown in figure 4.12.

4.4 Profilometers

4.4.1 Basic description

Profilometers use various techniques to gain info on the topography of a given surface. Within this thesis, two types of profilometers were used: optical and mechanical. Apart from their direct output of a geometry of a sample they are often used to indirectly gain other values. The most prominent usage that had effect on this work are die warpage scans and the transformation of die curvature to mechanical stress via the Stoney equation. Die warpage scans were performed by optical profilometry while mechanical profilometry was used for warpage scans of outer package surfaces.

4.4.2 Optical profilometer

Optical profilometers are non-contact surface profilers based on interferometry that result in measurements of a wide range of surface heights. The advantages of the optical based profilometer compared to the mechanical one is its wide scanning resolution adaptibility and capability of penetration through transparent materials. The measurements were performed with a Veeco Wyko NT3300 tool. The basic components of an optical profilometer are shown in *figure 4.13* while *figure 4.14* depicts the basic operation principle. The Veeco Wyko tool can be adjusted for usage of one of the two basic technologies:

- Phase-shifting interferometry (PSI)
- Vertical scanning interferometry (VSI)

In PSI, a white light is filtered and passed through an interferometer objective to the test surface. The interferometer beamsplitter reflects half of the incident beam to the reference surface wthin the interferometer. The beams reflected from the test surface and the reference surface recombine to form interference fringes. During the measurement, a piezoelectric transducer linearly moves the reference surface a small, known amount to cause a phase shift between the test and the reference beams. The system records the intensity of the resulting interference pattern at many different relative phase shifts and then converts the intensity to wavefront (phase) data by integrating the intensity data. The phase data are processed to remove phase ambiguities between adjacent pixels and the relative surface height can be calculated from the phase data as follows [111]:

$$h(x,y) = \frac{\lambda}{4\pi}\Delta\Phi(x,y) \tag{4.4}$$

The VSI represents a newer technique and shares the basic principles with PSI: light reflected from a reference mirror combines with light reflected from a sample to produce interference fringes, where the best-contrast fringe occurs at best focus. However, in VSI mode, the white light source is filtered with a neutral density filter, which preserves the short coherence length of the white light and the system measures the degree of fringe modulation, or coherence, instead of the phase of the interference fringes [111].

Optical mode	Height	Resolution
PSI	160 nm	1 nm
VSI	2 mm	3 nm

Table 4.2: Optical profilometer mode specifications

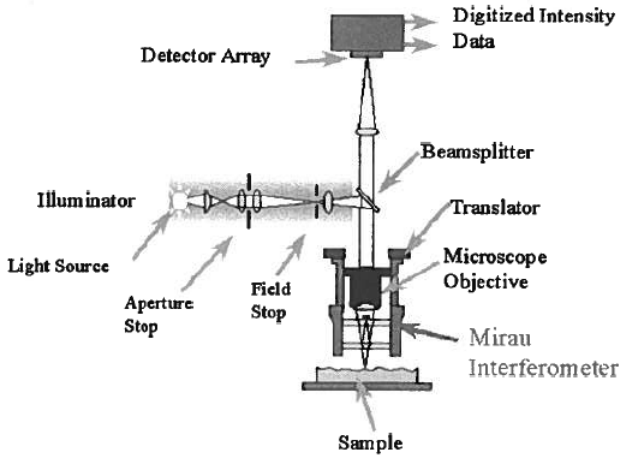


Figure 4.13: Basic components of an optical profilometer [111]

The slightly different operational modes of the Wyko system reflect more significantly in its applicability. In general, the PSI mode is finer and used for smoother areas. VSI is a rougher technique used to cover more prominent changes in surface heights. A basic performance regarding range and resolution of PSI and VSI techniques is summarized in *table 4.2*. Height refers to the largest vertical distance the profiler can accurately measure while resolution refers to the smallest vertical distance that can be accurately measured. Since the primary usage of the optical profilometer was warpage measurements across the die which deformation values are found in the μm range, VSI was standardly used for all measurements. *Figure 4.15* shows an example of a VSI surface scan of a ring type um range deformation, while *figure 4.16* shows its extracted curvature along a particular path. Such curvature graphs in terms of die warpage are later used for transformation to stress in silicon.

4.4.3 Mechanical profilometer

The mechanical profilometer serves the same purpose as the optical and has similar range and resolution performance. The main downside of the mechanical profilometer is its inherent operational principle of being able to investigate only surfaces it is in direct contact with while the optical profilometer analyzes surfaces through any transparent material, such as often used SiN die passivation. The mechanical profilometer

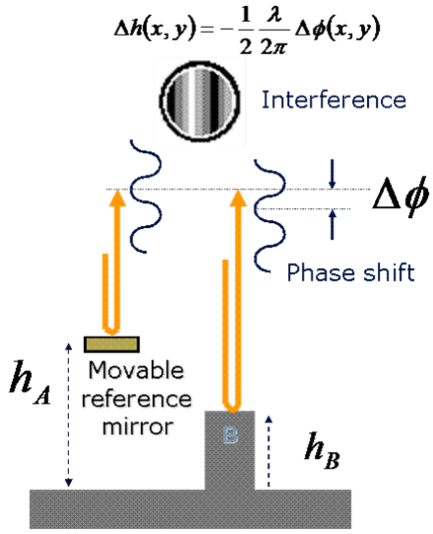


Figure 4.14: Basic operational principle of an optical profilometer[111]

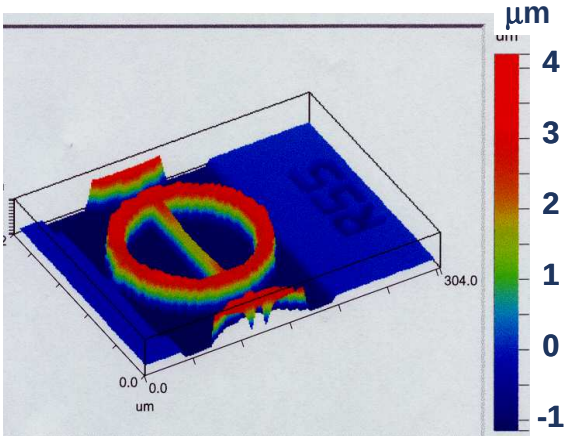


Figure 4.15: Surface scan of a ring type deformation[111]

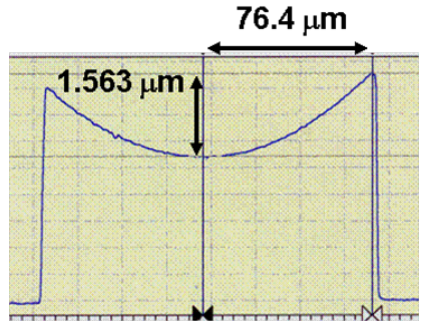


Figure 4.16: Extracted curvature from a specific position on the ring deformation [111]

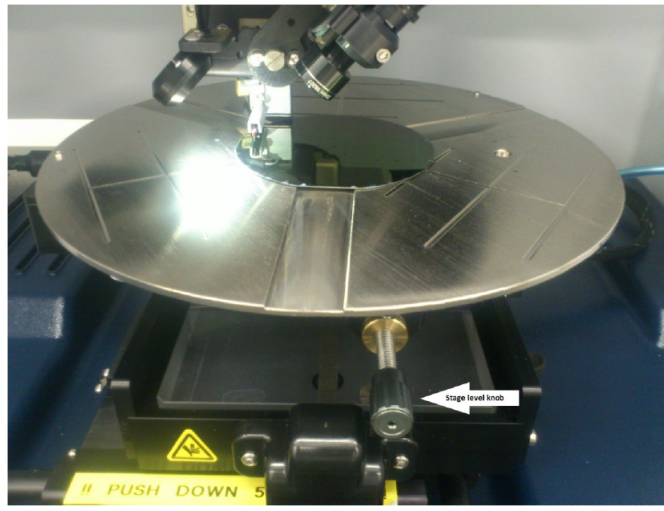


Figure 4.17: DEKTAK XT Stylus setup stationed in the lab [110]

Mechanical mode	Range	Resolution
DEKTAK	1 mm	1 Å

Table 4.3: DEKTAK XT Stylus specifications

is used within the thesis for profilometry of outward package surfaces.

The Bruker DEKTAK XT Stylus tool, displayed in *figure 4.*17, was used as a mechanical profilometer. Its basic properties are mentioned in *table 4.3*. Samples do not require any specific preparation.

4.4.4 Transformation of warpage to stress

The deformations obtained from profilometers can further be analyzed to establish a connection to stress. G.G. Stoney in 1909 derived a formula that describes stress in a two material stack plate system with uniform thicknesses [112] depending on the bending curvature of the system, presented in *eq. 4.5* and illustrated in *figure 4.*18.

$$\sigma_f = \frac{E_s h_s^2 r}{6 h_f (1 - \upsilon_s)} \tag{4.5}$$

Stoney's formulation states that the stress in a film, σ_f, on top of a substrate can be approximated using the material thicknesses of the film, h_f, and of the substrate, h_s, respectively, the substrate's Young's modulus E_s, Poisson's ratio υ_s and the curvature of the structure r. The equation is built on the following assumptions [113]:

1. The film and substrate thicknesses, h_f and h_s, are uniform, the film and substrate have the same radius r and $h_f \ll h_s \ll r$.

2. The film and the substrate are homogenous, isotropic and linear elastic

3. The stress states in the film are in-plane isotropic or equibiaxial (two equal stress components in any two mutually orthogonal in-plane directions) while the out-of-plane stress and shear stresses vanish.

4. The system's curvature components are equibiaxial (two equal direct curvatures) while the twist curvature vanishes in all directions

5. All surviving stress and curvature components are spatially constant over the system's surface

Stoney's equation is widely spread within engineering disciplines and extensively used for obtaining stress values in films from experimental curvature measurements. In its very simple form, it provides a very practical formulation necessitating variables which are usually well known allowing straightforward data collection. However, it has its limitations and is in essence a formulation bounded by the previously mentioned 5 points. Although in practice, all of these points are rarely satisfied, Stoney's equation still provides valuable first hand data and continues to be regularly used with its limitations and certain degree of error in mind. A fundamental property of curvature

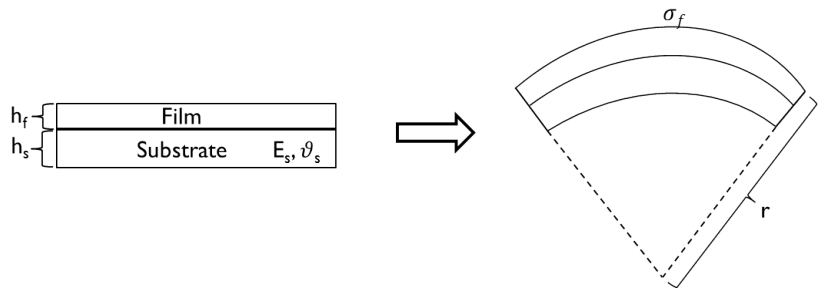


Figure 4.18: Warpage of a two material system and the accompanying values for stress calculation on the film surface

adds to the appeal of Stoney's equation. In case of a perfect circle, the curvature is constant throughout the curve and equals to the inverse of its radius, *eq* 4.9. Warped film-substrate structures are often approximated as circles in order to make use of this property.

$$\kappa = \frac{1}{r} \tag{4.6}$$

In reality, samples are not homogenous films and contain local structures. The validity of application of Stoney's equation is stretched when utilized to extract local stresses on a small area approximated as a circle type curvature. Further attempts have been made to address problems in point type usages of Stoney's equation and enhance its accuracy [113].

4.5 X-ray diffraction

4.5.1 Basic description

X-ray diffraction is a well known material inspection technique that finds it applications spanning through many domains from medicine to engineering. Today still evolving and adapting to usage in new technologies and fields, its origins date back to the beginning of the 20th century when von Laue, Friedrich and Knipping in 1911 discovered the phenomenon of diffracting X-rays in rock salt [114]. Following this discovery W.L. Bragg linked X-ray diffraction to evenly spaced sheets of atoms in crystal lattices introducing an equation mathemetically representing cases in which the diffractions will be constructive [115]. What is today known as Bragg's law and main formulation for constructive reflection of electromagnetic waves in crystals is presented in *eq. 4.7*

$$n\lambda = 2dsin\theta \tag{4.7}$$

where λ is the wavelength of the incident wave, θ its angle with the lattice sheets, d the interplanar spacing of the diffracting planes and n an integer representing the order of diffraction. In essence, Bragg's equation relates the distance between crystal

sheets and wavelength of the incident wave to the incident angle needed for constructive diffraction.

Within this thesis, the changing diffractive properties of X-rays are used to gain information on the stress and strain in Si encapsulated in a microelectronic package. A primary advantage of X-rays to other stress/strain inspection techniques such as mechanical/optical profilometers or Raman spectroscopy lies in its inherent ability of deep penetration through materials which allows a non-destructive analysis very appealing for investigation of mechanical effects in microelectronic packages. The results obtained with X-ray diffraction are a result of individual experiments during a PhD internship at GLOBALFOUNDRIES and a novel technique labeled X-ray diffraction 3D surface modeling (XRD-3DSM), developed and utilized by Dublin City University (DCU), where measurements were also performed.

4.5.2 XRD-3DSM technique

The XRD-3DSM is a novel technique for non-destructive analysis of strain/warpage inside completely encapsulated chips developed by DCU. Originally, the technique necessitated a synchrotron X-ray source [116,117] undermining widespread implementation, however latest efforts led to development of a non-destructive laboratory-based XRD tool [118]. The lab-based XRD-3DSM consists of a data reconstruction technique which monitors the 004 symmetric rocking curve (RC) full widhts at half maximum (FWHM) as a function of position across fully encapsulated packages [118]. Rocking curves, or ω-scans, are X-ray scans of specimens where the detector is placed at the center of the diffracted Bragg angle and the changing diffracted intensity is monitored while rotating, or "rocking", the specimen around the Bragg angle. The specifics of the lab-based XRD-3DSM technique are:

- during RC scans, the detector angle is fixed at twice the Bragg angle of the crystal planes of interest
- the incident X-ray beam is reduced to 250x250 μm by using horizontal and vertical slits in order to ensure a high spatial resolution
- a series of line scans are performed across the sample in steps down to 200 μm which are later converted to 3D warpage maps

A standard XRD-3DSM analysis starts by collection of rocking curves about the Bragg angle of Si 004 reflection. The sample is rotated around the [110] direction with respect to the Si surface normal [001] as indicated in *figure 4.*19. In this way, the FWHM of the RC is directly related to the lattice misorientation or in other words warpage of the Si die. *Figure 4.20 a)* shows an RC obtained from warped Si with a broadened FWHM while *figure 4.20 b)* depicts an RC of a flat Si die with a narrow peak around the Bragg angle. This comes from the fact that warpage creates a certain curvature of Si crystal planes resulting in a range of reflected angles for which the Bragg law is satisfied.

The rocking curves are obtained spatially throughout the Si surface and used to create an RC map. *Figure 4.21 a)* shows an RC map of a package Si die, scanned in 12 steps of 200 μm across its bottom edge. The map shows a variation of RC peak

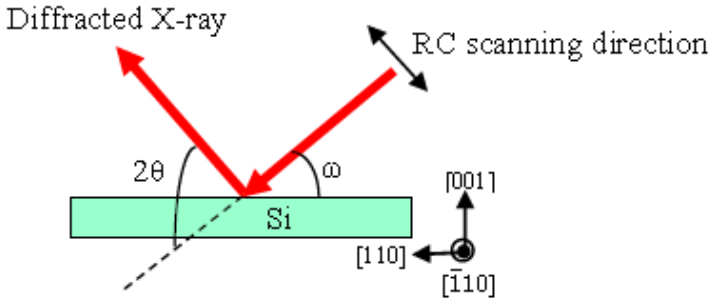


Figure 4.19: Basic XRD principle and rotation directions used in standard XRD scans [118]

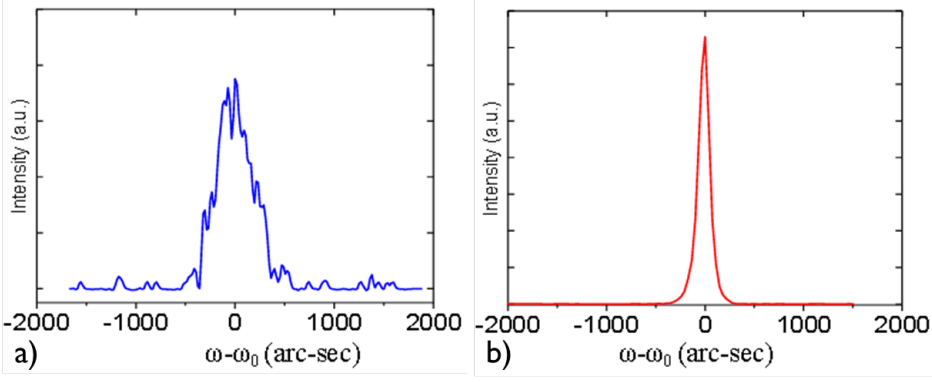


Figure 4.20: a) an RC of a warped Si die with a broadened FWHM and b) RC of a flat Si die with a narrow diffraction [118]

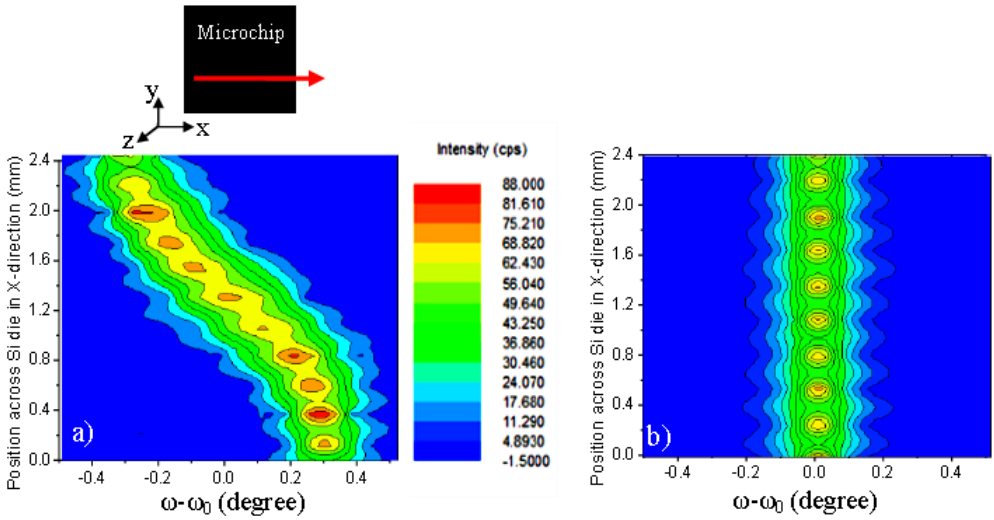


Figure 4.21: a) RC map of a warped Si die inside a package and b) an RC map of flat Si die [118]

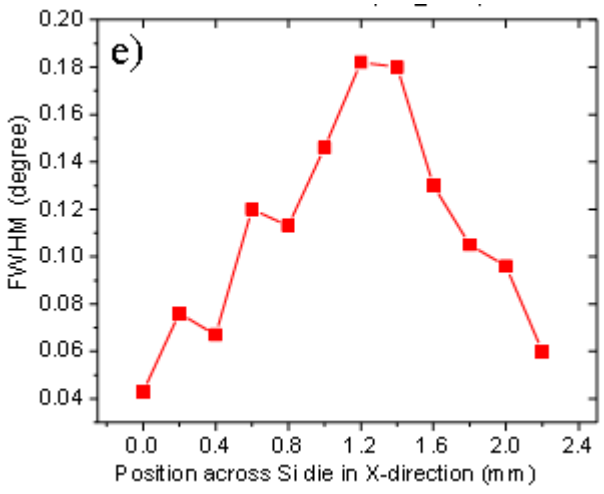


Figure 4.22: FWHM line extracted from an RC map scans [118]

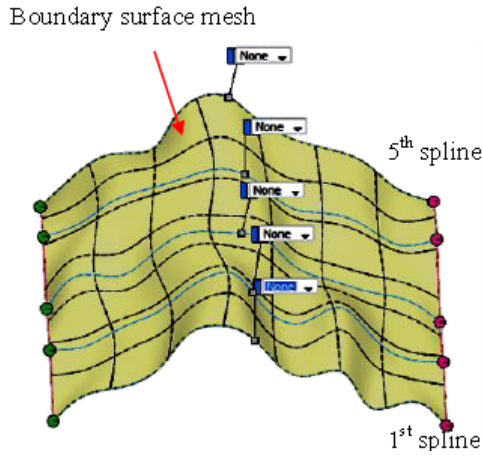


Figure 4.23: Surface mesh created from FWHM line scans depicting warpage over the whole Si surface [118]

positions ($\omega - \omega_0$) indicating presence of warpage. The diffraction angle on the left part of the Si die, where x<1.2, is larger than the angle in the center of the Si die (ω_0), while it decreases on the rigth part of the Si die. These particular shifts of the diffraction angle point out that the die is warped in a convex shape. *Figure 4.21 b)* presents an RC map of an unwarped Si die taken under the same conditions, where the diffraction angles remain the same throughout the die.

Therefore, the diffraction angle shifts visualized in RC maps can reveal the presence of warpage and its type, which is already useful in terms of non-destructively obtaining first indications of the Si die shape. However, members of DCU took a step further from here and went on to analyze the FWHM stating that just the diffraction angle shifts are not precise enough for a deep insight into the deformation and identifying highly distorted regions of the Si die. An approach was developed where each RC of the RC map was analyzed according to its FWHM and transformed into and FWHM line scan, as presented in *figure 4.*22.

Several RC scans are made in parallel to each other from the bottom edge of the Si die to its top edge and transferred to FWHM lines. Each FWHM line is represented with a spline curve forming a continous line and these spline curves in the end used to form a surface mesh, as shown in *figure 4.*23. Such a surface mesh represents a model created from direct measurements on the Si surface and provides detailed information on its lattice misorientations and warpage.

4.6 Test vehicles

In total 3 test chips, referred to also as test vehicles, were utilized in this thesis. Their names are ETNA, FUJI and PTCQ. These test chips were produced within the 3D integration program of imec. The ETNA, FUJI and PTCQ test chips were produced in

130 nm, 32 nm and 65 nm technology, respectively. The values 130 nm, 32 nm and 65 nm refer to the minimum feature sizes possible to process in that technology node. The standard definition for the value attributed to the technology node is the half-pitch of first level interconnect dense lines in dynamic random-access memory (DRAM) cells [119]. This however is not the smallest feature size processed on the chip. The smallest feature size on the chip relates to the gate length of a transistor. For instance, for a 130 nm technology node, the minimum processed half-pitch of first level interconnects is 130 nm, however the minimum transistor gate length is 90 nm [119].

The ETNA and FUJI chips are logic based test chips processed to investigate logic-on-logic and memory-on-logic behavior in 3D SICs. The PTCQ test chip is a mechanical stress dedicated test chip designed to monitor stress in Si after 3D IC stacking and 3D IC packaging. Transistors of interest were calibrated to stress on all 3 test chips. The ETNA test chip was stacked on a memory chip creating 2-die logic on DRAM prototypes. Since the stack is flipped when packaged, the packaged version is usually refered to as DRAM on logic. The FUJI test chips were stacked on each other creating 2-die logic-on-logic prototypes. The PTCQ test chips were single-die packaged in 2D technology, stacked on each other creating a 2-die 3D SIC and packaged in a 3D SIC package.

4.7 Summary

Usage of finite element analysis:

- predict mechanical stress in Si after 3D IC stacking and 2D and 3D packaging
- by utilizing test obtained transistor piezocoefficients, transfer simulated stress to transistor current shift
- compare with other stress extraction techniques
- upon finite element model validation with other stress extraction techniques, provide guidelines for stress reduction in Si

Usage of the delaminator:

- as a 4-pt bending tool to apply controlled in-plane stress during stress sensor calibration
- to apply controlled out-of-plane stress during stress sensor calibration

Usage of the nano-indenter:

- as an alternative to the delaminator for applying controlled out-of-plane stress during stress sensor calibration

Usage of profilometers:

- to determine the warpage of a Si die, 3D IC stack or package
- both a mechanical an optical profilometer is used, for double-check

- by obtaining the curvature of a die, using the Stoney equation, calculate stress in Si

Usage of X-ray diffraction:

- to determine warpage of packaged Si dies

Chapter 5

Impact of mechanical stress on the Front-End-of-Line

In this chapter, the impact of mechanical stress on FEOL devices is analyzed. The sensitivity of FEOL devices to stress is assessed from two points of view: to study the nature of stress impact on several types of FEOL devices through scaling technology nodes and to determine their suitability for CPI stress sensors. The motivation for assessing stress sensitivities of FEOL devices is introduced in section 5.1. Stress sensitivity of MOSFETs is analyzed in section 5.2. MOSFETs are calibrated to in-plane and out-of-plane stress in sections 5.2.1 and 5.2.2, respectively. Section 5.2.3 discusses in-plane stress extraction by means of MOSFETs. FinFETs and pseudo-Hall devices are calibrated to in-plane stress in sections 5.3.1 and 5.3.2, respectively. Additional studies on the in-plane stress sensitivity of MOSFETs are performed. Section 5.4 comprises tests of MOSFETs in a biaxial stress environment. Section 5.5 discusses impact of temperature on MOSFET sensitivity to stress. Section 5.6 discusses the need for stress calibration at high stresses and summarizes work on Si strength enhancement.

5.1 Impact of mechanical stress on scaling technology nodes

As discussed in sections 3.1 and 3.2, Si has a distinctive reaction to mechanical stress which can be measured as electrical current shift. This implies that all devices made on a Si basis exhibit a certain sensitivity to stress and that the FEOL may experience an impact from the mechanical stress generated in its surrounding. Stress in the FEOL has the ability to change the performance of FEOL devices and consequently circuits pushing them beyond their operating limits. A 3D IC environment introduces novel assembly steps and stress mechanisms, which can have a direct effect on Si.

The first experience of stress impact on FEOL in 3D IC, which became the basis of this thesis, was observed on the logic on DRAM prototype stack produced by imec in 2010. Figure 5.1 illustrates the 3D IC stack. A thin 25 μm Si logic die was stacked on a full thickness 700 μm memory die involving a thermo-compression bonding process using 13 μm high Cu-Sn microbumps and no-flow underfill material. The thermo-compression bonding was performed at 250 ºC. A 16x16 array of n-type MOSFET transistors with a channel width to length ratio of 800 nm by 600 nm processed in 130 nm technology were placed on the top side of the thin Si die, directly above the microbump. The projected microbump contour on the top Si surface is drawn as a yellow circle on the layout image of the FET array in figure 5.1. Another array of the same configuration was placed far away of the microbump, considered as a reference FET array. This means that the current shifts of the array above the microbump were referenced to the mean current value of the reference array transistors.

The drain currents of the reference and microbump FET arrays were measured after stack bonding and subsequent cooling to room temperature. Figure 5.2 a) presents the current shift results from the microbump FET array while figure 5.2 b) presents current shift results from the reference FET array, both referenced to the mean value of the reference FET array currents. A significant current shift of above 40% was observed on the transistors on the top Si side above the microbump while the reference array, far away from the microbump observed minor current scattering around its mean value. A relative change of a MOSFET drain current of 40% is far above any specifications and allowed MOSFET operating conditions. In certain cases, variations of up to 5% could be tolerated while variations below 1% would be desired.

Chapter 7 deals with the origin of the mechanism responsible for the current shift presented in figure 5.2 a) and its mitigation. This chapter analyzes the stress sensitivities of FEOL devices, primarily MOSFETs from two perspectives:

- To gain information on the nature of stress sensitivity of FEOL devices and its trends through scaling technology nodes
- Their applicability for usage as FEOL CPI stress sensors

Apart from MOSFETS, FinFETs and Pseudo Hall stress sensors were also analyzed.

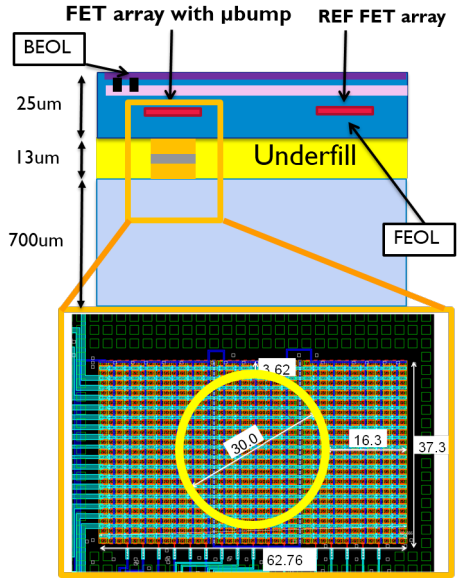


Figure 5.1: 16x16 array of 130nm technology node n-type MOSFETs placed on the top surface of the thinned Si die above a microbump and far away from it

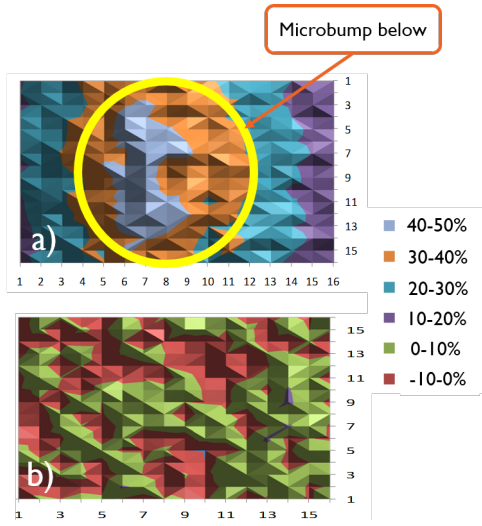


Figure 5.2: Current shift patterns from nFET arrays a) above the microbump position and b) from the reference array far away from the microbump. The nFET array above the microbump exhibits current shifts above 40% compared to the reference array.

5.2 Planar transistors

Planar transistors, namely MOSFETs are well known and widely used FEOL devices. The observed high current shifts from figure 5.2 due to stress mechanisms after 3D IC stacking point out the need for assessing stress behavior of MOSFETs in 3D IC circuits. The origin of the stress and amount of stress that lead to such high current shifts needs to be investigated. Some advantages of MOSFETs as potential stress sensing devices are:

- Familiarity - known operation and implementation in circuits, low cost
- Practicality - present in all circuitry today, accessible, minimal design effort if proposed as stress sensors
- Impact of temperature - although MOSFET properties are affected by temperature changes, circuit level temperature compensation for CPI evaluation studies is not needed as measurements are performed at room temperature, in temperature controlled environments and for technology development purposes, not field usage

Figure 5.3 presents a typical cross section of a MOSFET. *S* reprents the source contact, *G* the gate contact, *D* the drain contact, *B* the bulk contact and voltages V_{GS} and V_{DS} the gate to source and drain to source voltages, respectively. For n-type MOSFETs, the source and drain are heavily doped n-type islands imbursed in a p-type substrate. For p-type MOSFETs, the source and drain are heavily doped p-type islands imbursed in a n-type substrate. For simplicity, the depletion regions between p-n contacts are ommited. The conductive channel between source and drain contacts is created below the gate and oxide layer. The drain to source current, I_D, exhibits a shift in value after application of external mechanical load.

Eq. 5.1 [120] presents the drain to source current, I_D, in linear region while eq. 5.2 [120] presents the drain to source current in saturation region. The drain to source current in MOSFETs, I_D, is directly proportional to the current carrier mobility, μ. A shift in mobility will directly cause a transistor drain to source current shift.

$$I_D = \mu C_{ox} \frac{W}{L} \left[(V_{GS} - V_{th}) V_{DS} - \frac{V_{DS}^2}{2} \right] \tag{5.1}$$

$$I_D = \frac{1}{2} \mu C_{ox} \frac{W}{L} (V_{GS} - V_{th})^2 \left[1 + \lambda (V_{DS} - V_{DSsat}) \right] \tag{5.2}$$

W and L stand for the transistor channel width and length, respectively, and C_{ox} represents the capacitance of the gate to channel oxide per unit area. V_{GS} and V_{DS} stand for the gate to source and drain to source voltage bias and V_{th} stands for the transistor threshold voltage. An additional correction factor impacting the slope of the saturation current is given by λ. λ is a process-technology parameter that, for a given process, is inversely proportional to the channel length [120].

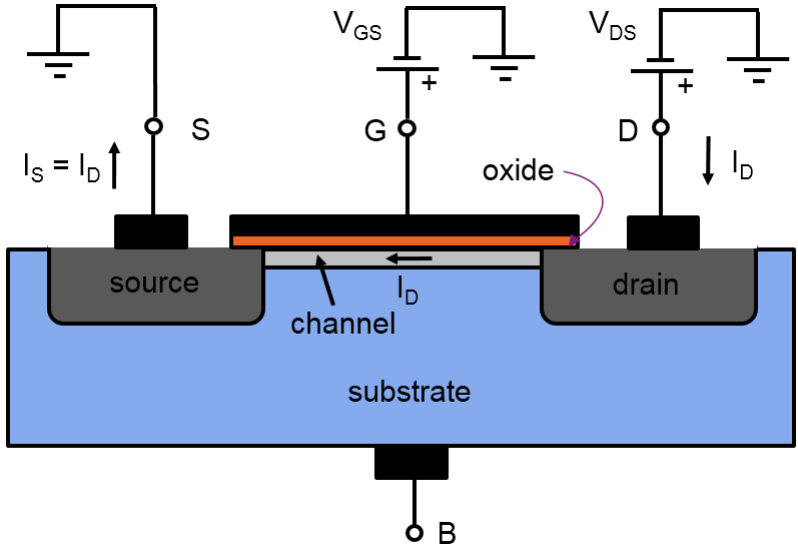


Figure 5.3: A typical MOSFET cross-section with external contacts

Figure 5.4 shows the typical I_D-V_{GS} curve of a n-type MOSFET while figure 5.5 shows a typical I_D-V_{DS} curve of a n-type MOSFET. Typical p-type MOSFET I-V curves are the same as n-type MOSFET curves with the difference that applied voltages and drain current are of opposite sign. A channel is created when the applied gate voltage, V_{GS}, surpasses a certain value, the transistor threshold voltage, V_{th}, attracting enough carriers from the substrate bulk to the substrate surface. If the drain and source voltage, V_{DS}, is biased to create a lateral electric field in the channel and the gate voltage, V_{GS}, is higher than the treshold voltage, V_{th}, current flow is enabled as indicated in figure 5.4. If the gate voltage, V_{GS}, higher than the treshold voltage, V_{th}, is kept constant and drain to source voltage, V_{DS}, is varied, two distinct MOSFET operation regimes can be distinguished. For $V_{DS} < V_{GS} - V_{th}$ the transistor operates in the linear regimes and for higher V_{DS} values where $V_{DS} \geq V_{GS} - V_{th}$ the transistor operates in the saturation regime.

If a MOSFET's current shift under mechanical stress is considered a consequence of the change of Si channel mobility then its behavior can be interpreted through the piezoresistance effect (section 3.1). This implies that if a MOSFET under the same biasing conditions is considered as a stable resistor, eq. 3.5 can be applied connecting mobility shift, resistivity shift and current shift. In essence, the mobility shift in the drain-channel-source region is reflected into resistivity shift (eq. 3.3) which is measured as drain to source current shift. The relative resistivity change in eq. 3.7 can then be replaced by the relative current shift leading to eq. 3.10 and 3.11, repeated in general form below:

$$\frac{\Delta I_D}{I_D} = \pi_l \sigma_l + \pi_t \sigma_t + \pi_v \sigma_v \tag{5.3}$$

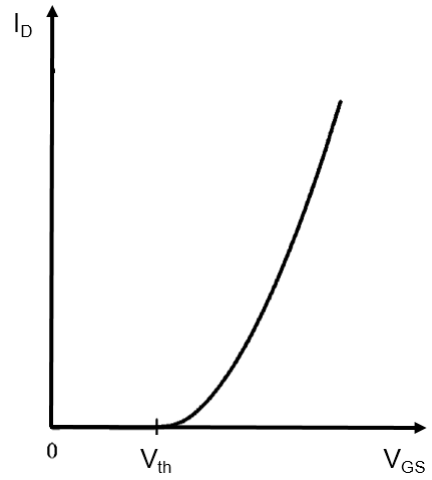


Figure 5.4: A typical n-type MOSFET I_D-V_{GS} curve

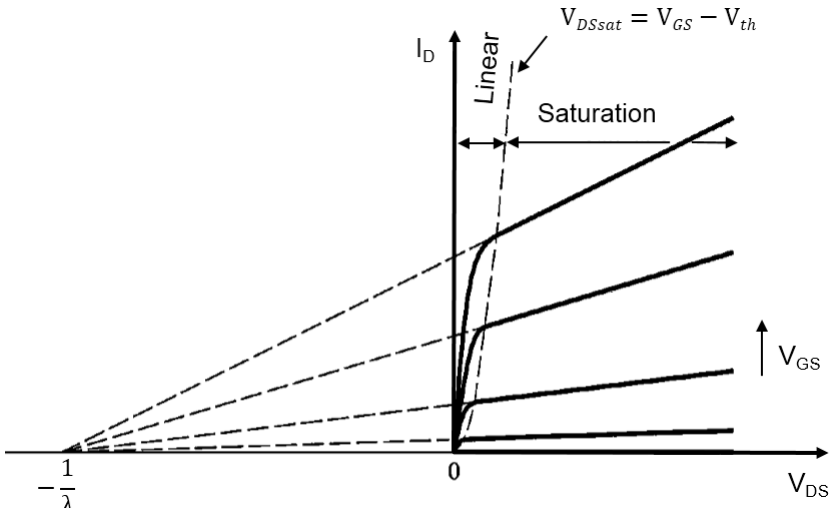


Figure 5.5: Typical n-type MOSFET I_D-V_{DS} curves

where I_D is the transistor drain current, π_l, π_t and π_v the transistor longitudinal, transverse and vertical piezocoefficients, respectively, and σ_l, σ_t and σ_v the corresponding normal stresses acting on the transistor. Eq. 5.3 in this form does not take into account the impact of:

- shear stresses
- threshold voltage
- temperature

Looking at bulk Si, if the relative resistance shift in eq. 2.3 is replaced by current shift, shear stress $\sigma_{12}^{'}$ has no contribution to the current shift when $\phi = 0$, or in other words shear stress has no impact in [110] direction and 90° shifts from this direction if the current is parallel or perpendicular to these directions. The standard reference frame in microelectronics processing involves transistors produced in a [110]/[-110] reference frame with a [001] directed Si surface.

The impact of stress on transistor threshold voltage can not be covered by the piezoresistance model. Electrical measurements have been performed to extract the impact of stress on the threshold voltage on particular transistors and are presented at the end of section 5.2.1. No impact of stress on the threshold voltage was observed. Section 5.5 presents tests related to piezocoefficient changes at higher temperatures. Furthermore, if solely the MOSFET channel resistance shift needs to be extracted, the drain and source resistance need to be deducted from the overall transistor resistance. This is however of no particular interest to this thesis, as the important measured value is the drain current shift representing the transistor reaction to stress on device level as a whole, rather than channel level.

5.2.1 In-plane stress sensitivity

A series of MOSFETs throughout 3 IMEC processed technology nodes were calibrated to in-plane stress in [110] and [-110] direction according to the procedure explained in sections 3.2.1 and 3.2.2. The summarized results are presented in table 5.1 along with the bias conditions summarized in table 5.2. Piezocoefficients obtained in the [110] direction are referred to as longitudinal, while piezocoefficients obtained in [-110] direction are refered to as transverse. Both n-type and p-type, short and long channel MOSFETs from 130nm, 65nm and 32nm technology nodes were calibrated in their saturation and linear region. Smith's piezocoefficients [93] obtained for low-doped bulk Si are placed in the last column for comparison. The number of calibrated transistors varied from test vehicle to test vehicle, depending on sample size availability and allowed time. In general the number of transistor per each type tested ranged from ten devices to more than a hundred devices.

The sensitivity to stress was extracted for each individual transistor and then the median value of a group of sensitivities from all of the transistors of the same type was taken as the reference sensitivity of that tranistor in that particular technology. In order to provide an idea on the scattering of the obtained transistor sensitivities, table 5.3 summarizes the maximum standard deviation of the obtained transistor piezocoefficients for each technology node. This means that the values in table 5.3 represent the

Technology node	FET type	Width/Length (nm/nm)	Piezocoeff. type	Piezocoeff. Saturation (ppm/MPa)	Piezocoeff. Linear (ppm/MPa)	Smith piezocoeff. (ppm/Mpa)
130 nm (ETNA)	N, long	800/600	longitudinal	-331	-	-312
	N, short	800/130	longitudinal	-252	-	-312
	N, long	800/600	transverse	-168	-	-176
	N, short	800/130	transverse	-101	-	-176
65 nm (3D 65)	N, long	5000/10000	longitudinal	-408	-	-312
	N, short	500/500	longitudinal	-180	-	-312
	P, long	5000/10000	longitudinal	690	-	718
	P, short	500/500	longitudinal	585	-	718
32 nm (FUJI)	N, long	500/900	longitudinal	-215	-248	-312
	N, short	500/70	longitudinal	-100	-120	-312
	N, long	500/900	transverse	-135	-155	-176
	N, short	500/70	transverse	-51	-88	-176
	P, long	500/900	longitudinal	406	457	718
	P, short	500/70	longitudinal	73	110	718
	P, long	500/900	transverse	-375	-440	-690
	P, short	500/70	transverse	-165	-249	-690

Table 5.1: In-plane stress piezocoefficients obtained with the adapted 4-pt bending setup for various MOSFETs processed on IMEC test vehicles through 3 technology nodes. Smith piezocoefficents are present for comparison to low doped bulk Si.

Technology node	FET type	Operating regime	Source (V)	Drain (V)	Gate (V)	Bulk (V)	Power supply (V)
130 nm (ETNA)	N	Saturation	0	0.9	0.9	0	-
65 nm (3D 65)	N	Saturation	0	1	1	0	-
	P	Saturation	1	0	0	1	-
32 nm (FUJI)	N	Saturation	0	1.2	1.2	0	1.2
	N	Linear	0	0.05	1.2	0	1.2
	P	Saturation	1.2	0	0	1.2	1.2
	P	Linear	1.2	1.15	0	1.2	1.2

Table 5.2: Bias conditions for MOSFETs calibrated to in-plane stress from table 5.1

Technology node	Test vehicle	Max. standard deviation [MPa]
130 nm	ETNA	$\pm$15
65 nm	3D 65	$\pm$10
32 nm	FUJI long	$\pm$18
	FUJI short	$\pm$13

Table 5.3: Maximum standard deviation of the obtained piezocoefficients observed on particular transistor types for each technology nodes

highest standard deviation observed on a group of devices of the same type within the same technology node, acting as the worst case scenario. For all technology nodes and all measured transistors, the maximum observed standard deviation of the stress sensitivity was between 10 MPa and 20 MPa. Figures 5.6, 5.7, 5.8 and 5.9 present examples of calibration results from all the technology nodes and device types, in accordance with table 5.1.

The following observations can be made from the obtained sensitivities in table 5.1 and corresponding figures 5.6 to 5.9:

- MOSFETs are sensitive to in-plane normal stress
- The obtained piezocoefficients differ from Smith's piezocoefficients
- The obtained piezocoefficients differ from technology node to technology node
- Longer channel devices exhibit higher sensitivity than shorter channel devices
- MOSFETs in linear regime are more sensitive to stress than in saturation regime

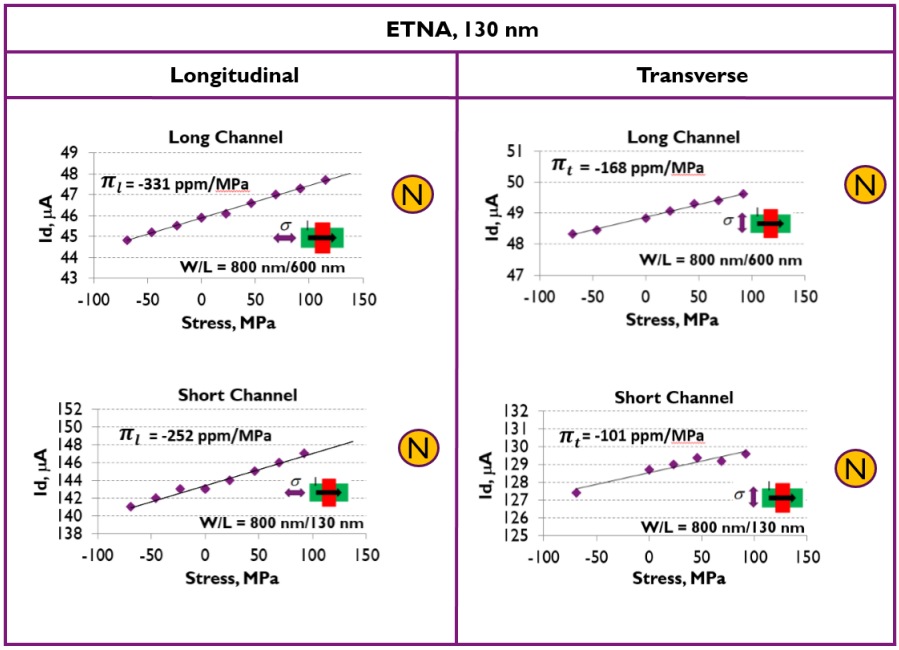


Figure 5.6: Examples of current shift calibration results to in-plane stress for all device types evaluated on the ETNA 130nm technology node test vehicle

Bulk planar transistors such as MOSFETs use the electrical properties of bulk Si for modulation of device current flow. Electrical changes in Si are directly transferred to device operation. MOSFETs are known to be sensitive, based on tests performed on older technology nodes involving mostly large channels [85-88]. Contemporary transistors used in table 5.1, from the 130 nm down to the 32 nm node, confirmed their sensitivity to in-plane stress. A piezocoefficent value of 200 to 300 ppm/MPa points to 2 to 3% of current shift under a uniaxial load of 100 MPa. Current shift would increase by approximately another 1-2% in a biaxial state of 100 MPa reaching current shifts of up to 5% when the transverse in-plane stress impact is included. For instance, 2D chip overmolding processes often induce in-plane stress between 100 and 200 MPa, which according to our scenario potentially raises the current 5-10% .

3D SIC introduces new complex assembly steps which can substantially increase the ones in a 2D environment, reaching several hundreds of MPa. Having said that, there are two ways we can look at the obtained stress sensitivities of MOSFETs:

- as devices which are sensitive to stress requiring action to ensure stable operation
- as stress sensors able to measure the stress in Si

In the first case, the in-plane piezocoefficients give us an insight into the behavior of the transistor in a biaxial in-plane stress environment. If stress present in Si is known up-front, transistors could be placed at sufficient distance from high stress areas. Exactly

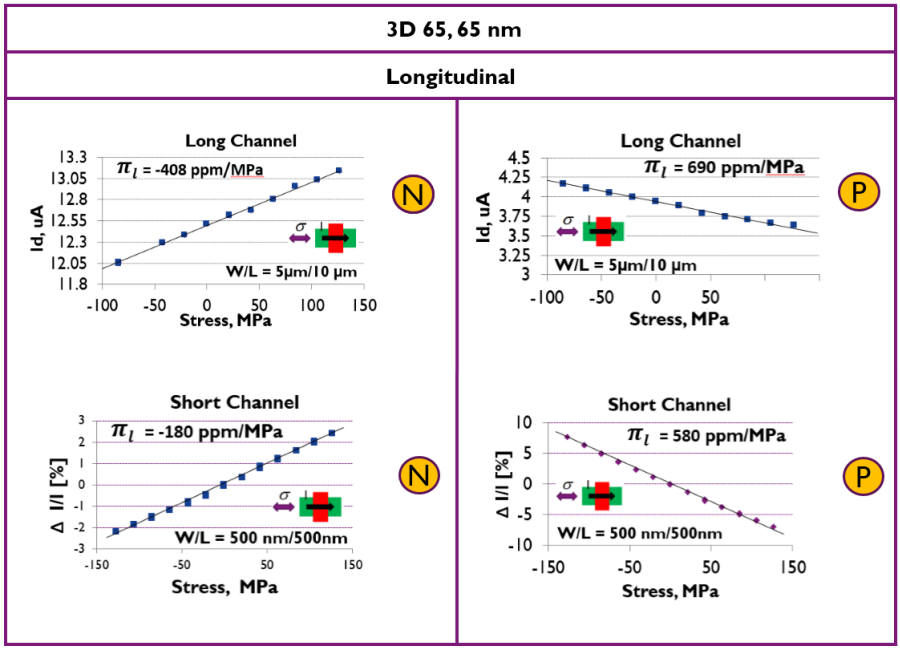


Figure 5.7: Examples of current shift calibration results to in-plane stress for all device types evaluated on the 3D 65 test vehicle, 65nm technology node

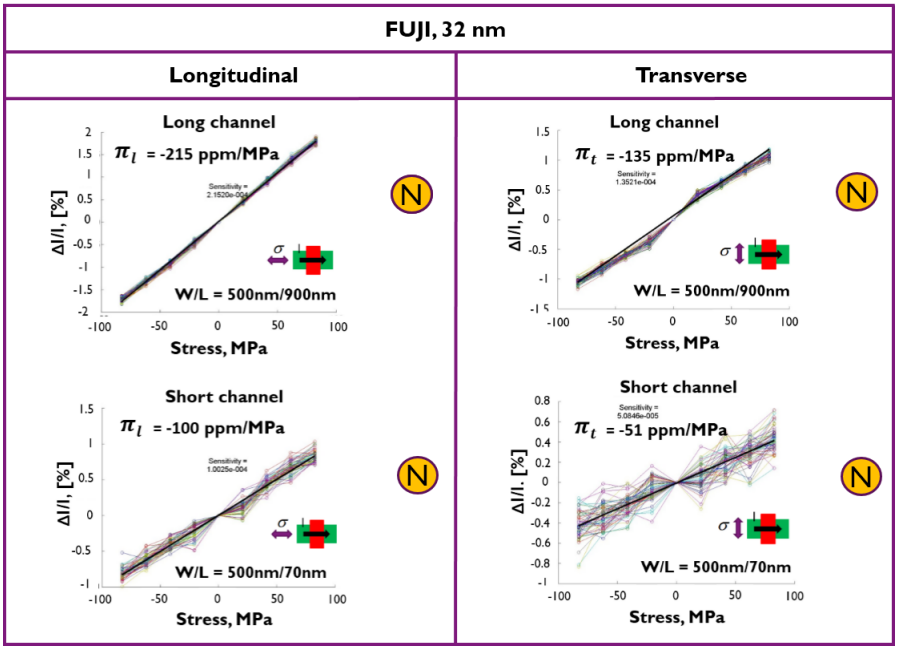


Figure 5.8: Examples of current shift calibration results to in-plane stress for n-type devices evaluated on the FUJI 32nm technology node test vehicle

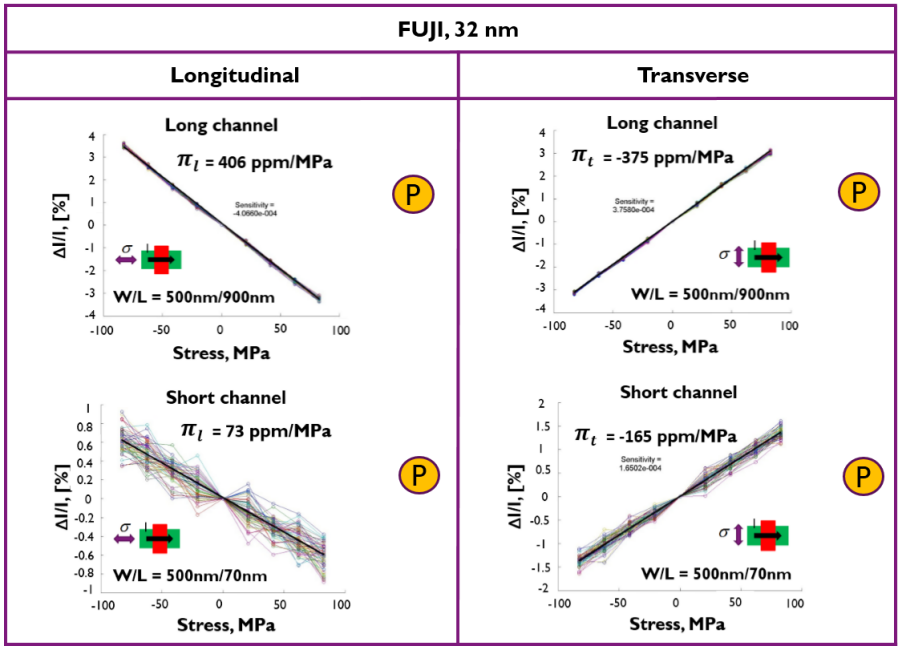


Figure 5.9: Examples of current shift calibration results to in-plane stress for p-type devices evaluated on the FUJI 32nm technology node test vehicle

how far would be defined by keep-out-zones in IC layouts based on the obtained transistor stress sensitivities, as in table 5.1. In literature, calibration of FEOL devices to stress is rarely done. According to findings from table 5.1, transistor sensitivities can substantially differ from piezocoefficients obtained by Smith in 1954 on low doped bulk Si. Furthermore, a great variety of sensitivities can be present through technology nodes and devices within the same technology node. Utilizing Smith's piezocoefficients for Si based FEOL devices can introduce a significant error in current shift or stress calculation. Therefore, in terms of evaluating impact of stress on FEOL devices, it is crucial to obtain device sensitivities through direct calibration. This is valid not just for in-plane stress components. Impact of out-of-plane stress will be discussed in section 5.2.2.

In order to extract stress values in Si upfront and understand stress generation in Si originating from various assembly processes, Si based transistors acting as stress sensors seem to be a logical approach. In that sense, some MOSFETs from table 5.1 prove to be sensitive enough to be considered as stress sensors. In this case, the higher the sensitivity to stress, the better.

In every technology node, shorter channel transistors exhibit lower sensitivity to stress than longer channel transistors. Apart from known short channel effects or other effects that might affect carrier mobility, a shorter channel's resistance shift from the viewpoint of source to drain resistance will have a lower relative impact on the overall drain to source resistance than a larger channel. In other words, even if the resistance shift in the channel of two transistors with different channel lengths is the same, the overall device sensitivity will be different due to the different relative impact of the channel resistance in the whole source to drain resistance path. As the source and drain are usually heavily doped, their mobility shifts are negligible, however not their resistances. Intrinsic sensitivity of the channel can be calculated if the source and drain resistance are taken into account. Within this thesis, calculating intrinsic sensitivity is not necessary for evaluating the device's sensitivity whether in terms of stress sensors or stress impact evaluation as the real response of the device is a sum of all its elements representing the device as one whole, not just its channel.

Furthermore, the linear region has a slight sensitivity advantage to the saturation regime. The mobility of carriers is directly related to the electric field. In saturation regime under higher electric fields, a rise in electrical field will have a limited impact on the mobility and electrical current. Under lower electric fields, as in the linear regime, the mobility is not saturated and hypothetically has more room for shift under mechanical load.

The piezoresitance model takes into account mobility shifts in Si. A MOSFETs current is not based only on the applied drain to source voltage, but the behavior of the gate-oxide-semiconductor structure. A transistor is open and conducting current only if the voltage applied to the gate surpasses the transistor's threshold voltage V_{th}. A shift of V_{th} under mechanical load would have a direct impact on the transistor's drain to source current.

For calibrated 130nm, 65nm and 32nm technology node devices, the current shift response to in-plane stress was linear on all devices. The assumption is that either the stress does not impact the transistor's V_{th}, or the stress impact on V_{th} causes also a linear reponse of the drain to source current shift superpositioning itself to the drain to source

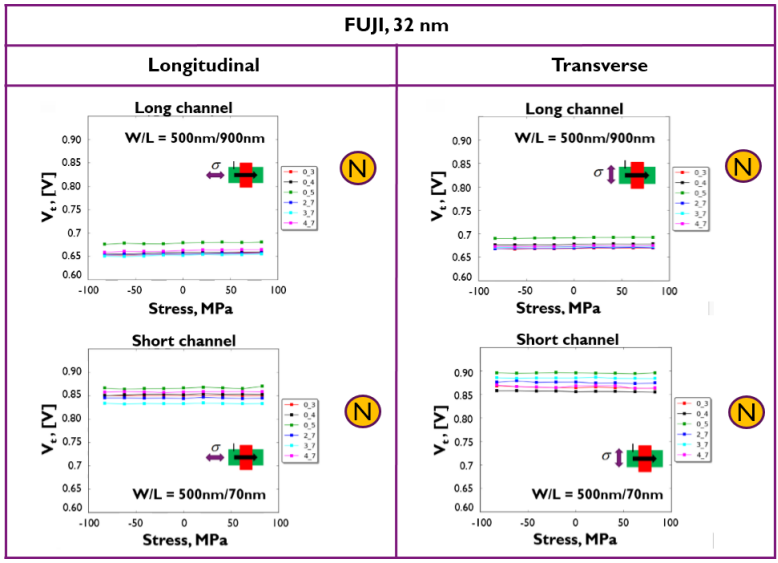


Figure 5.10: Examples of threshold voltage calibration results to in-plane stress for n-type devices evaluated on the FUJI 32nm technology node test vehicle

current shift caused by the channel mobility shift.

The V_{th} was extracted under mechanical load for 32nm technology node transistors. Figures 5.10 and 5.11 summarize the results. No threshold voltage shift was observed on any devices. This furthermore supports the usage of the basic piezoresistance model for MOSFETs where current shift in response to stress is attributed to the bulk Si mobility shifts.

A general advantage of transistor stress calibration is collecting all the impacts to current shift, including the ones which might not be apparent and mentioned here. A transistor sensitivity value can arguably be discussed in terms of its channel mobility shift, threshold voltage shift, direction of applied stress and so on, however the final calibration curve provides the real response of the transistor, with all possible effects combined. If this curve is linear, which is the case on all tested transistors, from 130nm node down to 32nm node, a sensitivity value can be extracted which clearly describes its behavior in a stress environment. From a scientific point of view, further investigation on the contributions to current shift can be made. From an engineering point of view, the final sensitivity value gives enough information to characterize the devices response to stress, whether in terms of benchmarking the particular device of the technology node or consideration of devices for stress sensors. In this section, transistor calibration to in-plane stress was presented and contributions of shear stress discussed. The following chapter presents results on MOSFET calibration to out-of-plane stress.

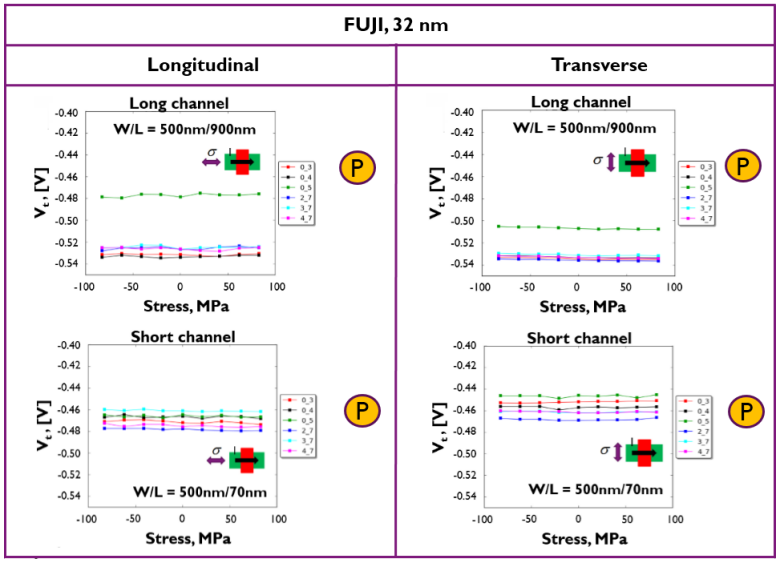


Figure 5.11: Examples of threshold voltage calibration results to in-plane stress for p-type devices evaluated on the FUJI 32nm technology node test vehicle

5.2.2 Out-of-plane stress sensitivity

According to Smith's stress calibrations [83,93], in [001] direction, the out-of-plane direction of standard Si wafers, n-type Si is highly sensitive to stress while p-type Si exhibits very low sensitivity. Calibration of transistors to out-of-plane stress was performed on 65nm technology n-type transistors of the PTCQ test vehicle. PTCQ is a stress dedicated test chip developed in-house based on all prior learnings from test vehicles ETNA, 3D 65 and FUJI. The design of the PTCQ test chip and all results will be presented in chapters 7 and 8. This section includes transistors from the PTCQ test chip only in terms of out-of-plane transistor calibration, as a logical continuation of chapter 5.

Long n-type transistors with channel width to length ratio of 4x4.4 μm were calibrated using the delaminator and nano-indenter tool. The procedures are explained in section 4.2.3.1 for the delaminator and 4.3 for the nano indenter. Figure 4.8 depicts the delaminator out-of-plane procedure using a Si cube while figure 4.12 depicts the spherical tip used in the nano-indenter procedure. Figure 5.12 a) shows the IC layout of the transistors grouped in a 7x8 array. The yellow circle depicts the position of a 50 μm diameter Cu pillar that was positioned on the BEOL above the transistor array.

In case of calibration with the delaminator, the pressed Si cube distributed the delaminator applied force over 196 Cu pillars lying beneath. The acting out-of-plane stress was calculated as the delaminator force over the area of the 196 Cu pillars. In case of the calibration with the nano-indenter, one Cu pillar was pressed with a 250 μm spherical tip. Figure 5.12 b) presents 2 out-of-plane stress calibration curves obtained with the nano-indenter and 1 out-of-plane stress calibration curve obtained with the

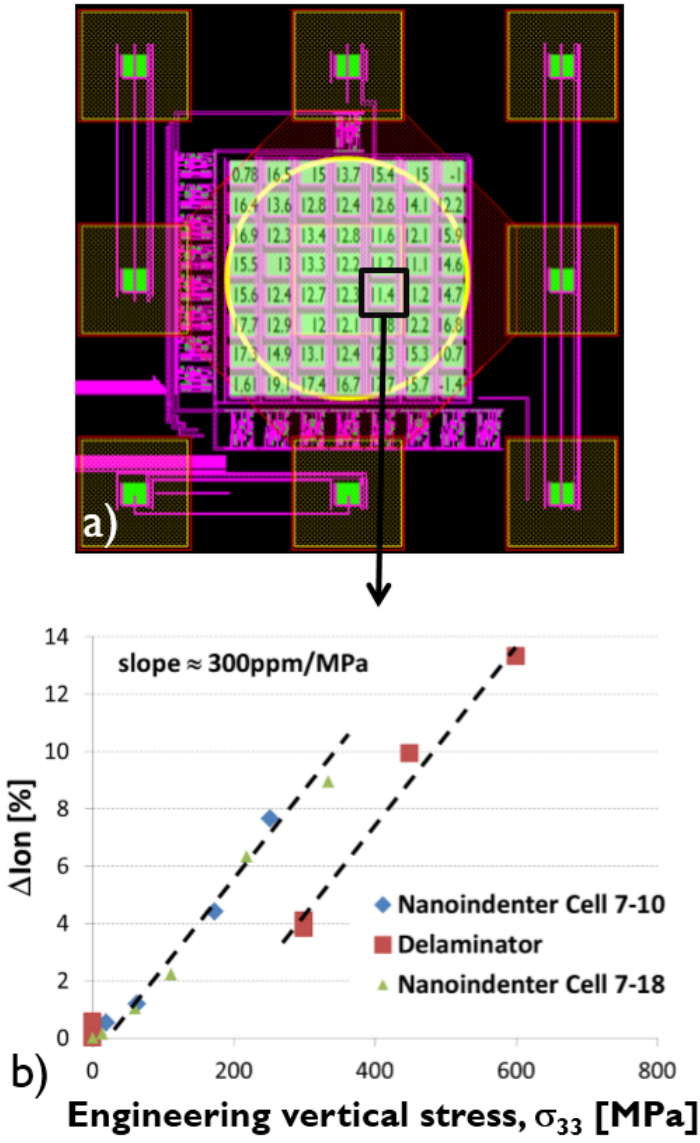


Figure 5.12: Current shift calibration to out-of-plane stress: a) IC layout of the transistors grouped in a 7x7 array with the yellow circle indicating the position of the Cu pillar above the array, b) the results of the out-of-plane calibration using the delaminator and the nano-indenter

Technology node	FET type	Operating regime	Source (V)	Drain (V)	Gate (V)	Bulk (V)	Power supply (V)
65 nm (PTCQ)	N	Saturation	0.05	0.95	1	0	1

Table 5.4: Bias conditions for MOSFETs calibrated to out-of-plane stress from figure 5.12

delaminator for the n-type transistor. Table 5.4 summarizes the bias conditions for the MOSFETs during calibration to out-of-plane stress.

A linear dependence of the current shift and externally applied out-of-plane stress is observed. A jump is observed from the first to the second measurement point when using the delaminator. It is believed that this is due to the varying heights of the 196 Cu pillars under the Si cube. When the Si cube is pressed down on the Cu pillars, due to the Cu pillar height variation, not all Cu pillars are initially touched. Some Cu pillars need to most likely plastically deform to allow contact to all other, lower positioned Cu pillars. For this particular delaminator obtained measurement in figure 5.12 b), it is likely that the second measurement point does not correspond to approximately 300 MPa that can be read from the x-axis, but a lower value, consistent with the nano-indenter obtained values. However, the further delaminator obtained slope obtained from points including and following the second measurement point, give a similar slope to the nano-indenter obtained slope.

A mean piezocoefficient value of 300 ppm/MPa is observed for n-type transistors. This confirms that the impact of out-of-plane stress on MOSFETs should not be neglected and that the out-of-plane stress calibration should accompany the in-plane stress calibration.

A discussion can be made in which form the out-of-plane stress applied on the Cu pillar reaches the bulk Si transistor. If the BEOL would be ideally considered consisting of homogenous layers, the stress to Si would be the same as applied to the Cu pillar, therefore known through the force monitored with the delaminator or nanoindenter. In practice, each BEOL layer consists of a combination of Cu vias surrounded by low-k material and silicon oxide. The Cu vias are significantly stiffer than the surrounding low-k material and with an order of magnitude lower diameter than the Cu pillar, will attract higher stress. If a Cu via surrounded by low-k material was directly connected to the transistor gate, this might cause high local stress in Si, higher than observed on the Cu pillar above. However, the gap between the first metal layer and active Si is usually filled with silicon oxide which is significantly stiffer than low-k, comparable to Si and Cu. Therefore, the chances of having excess local stress in Si due to Cu vias or in general due to variations in BEOL design are minimized, although not excluded for certain.

Furthermore, out-of-plane stress application in such a manner over a Cu pillar can cause also in-plane stress components in Si. The sensitivity value obtained from this calibration would then be a sensitivity to a combination of out-of-plane and in-plane stress. If the amount of in-plane stress caused during the calibration is known or at least

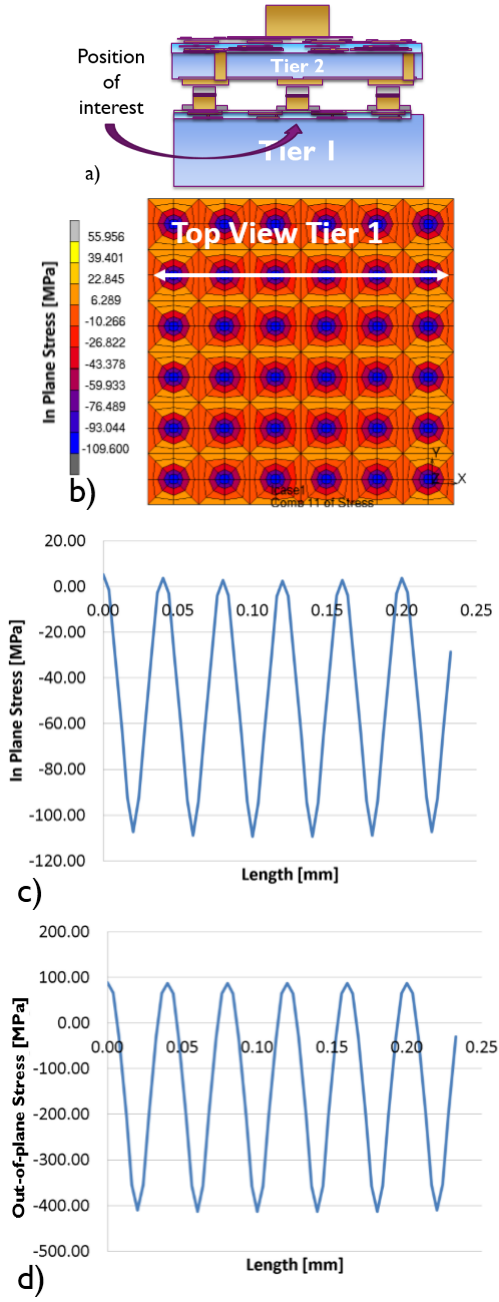


Figure 5.13: Example of out-of-plane stress in a 3D SIC stack: a) position of interest based on which a finite element model was built, b) the FEM obtained in-plane stress pattern observed in Si with circular stress regions corresponding to microbump positions, c) the in-plane stress plotted in Si below several microbumps and d) the out-of-plane stress plotted in Si below several microbumps, path indicated in b).

Out-of-plane sensitivity	N-type piezocoeff. [ppm/MPa]	P-type piezocoeff. [ppm/MPa]
From Smith	534	-11
Initial calibration	300	-
Corrected calibration	433	-

Table 5.5: Summarized out-of-plane piezocoefficient values for n-type devices in [001] direction, obtained from the PTCQ 65 nm technology node test vehicle

the ratio of the caused in-plane stress and externally introduced out-of-plane stress is known, additional corrections to the out-of-plane piezocoefficient can be made. For this purpose, a finite element model dealing with underfill-microbump induced stress in Si discussed in section 7, will here be used to further discuss the out-of-plane piezocoefficient correction. Detailed explanation on the origin of the stress within the 3D stack will be given in chapter 7.

Figure 5.13 a) presents an illustration of a 3D SIC after which the finite element model was built. After thermocompression bonding of the dies over a Cu-Sn microbump and cooling down to room temperature, stress is observed in the Si below a microbump, the position of which is indicated with an arrow. Figure 5.13 b) represents the in-plane stress field of the top side of the die labeled Tier 1,below the microbumps. Figures 5.13 c) and d) provide graphs with detailed curves of the in-plane and out-of-plane stress along the path indicated in figure 5.13 b). As will be shown in section 7, the two independent in-plane stress components can be considered equal in the Si surrounding the microbump. The 3D SIC in figure 5.13 provides a similar setting as the one observed during out-of-plane stress calibration. A Cu-Sn microbump is pressed in the out-of-plane direction on the BEOL and FEOL causing local stress in Si beneath it. From the simulations it follows that, along with out-of-plane stress, in-plane stress is observed, approximately 3.6 times lower.

Eq. 3.10 connecting the three independent normal stress components with current shift, in case of equal two in-plane stress components, can be rewritten as

$$\pi_v = \frac{\Delta I}{I} \cdot \frac{1}{\sigma_v} - R(\pi_l + \pi_t) \tag{5.4}$$

where R stands for the ratio of the in-plane stress and out-of-plane stress, σ_l/σ_v. The first term in the equation is actually our directly obtained out-of-plane piezocoefficient, 300 ppm/MPa, which will now be modified with the ratio R and the in-plane piezocoefficients. For R of 3.6 and in-plane piezocoefficients of -306 ppm/MPa and -180 ppm/MPa obtained from in-plane calibrations of the particular devices on the PTCQ testchip, the piezocoefficient value π_v now changes to 433 ppm/MPa. This value is closer to Smith's out-of-plane piezocoefficient value for n-type bulk Si. The corrected out-of-plane piezocoefficient value should be utilized as the more realistic

value linking out-of-plane stress to current shift. Table 5.5 summarizes the out-of-plane piezocoefficient values for n-type Si in [001] direction, mentioned in this section.

5.2.3 Extracting in-plane stress

Ideally, a set of stress sensors would strive to extract all individual stress components. In practice, only a portion of these requirements can be satisfied. Sections 5.2.1 and 5.2.2 displayed MOSFET sensitivity to in-plane stress in [110] and [-110] direction and out-of-plane stress in [001] direction. According to eq. 3.7 and eq. 3.5 linking stress components to current shift, all three normal stresses add up to the final current shift value. In terms of stress extraction, the same measured transistor current shift is linked over previously obtained piezocoefficients to 3 independent normal stress components, representing one equation with 3 unknowns.

Arranging a set of MOSFETs in a particular fashion can lead to in-plane stress extraction with certain constraints. In-plane stress in [110] and [-110] direction can be extracted from an approximately two transistor sized area by placing one transistor with current flow in [110] direction and the other next to it with current flow in [-110] direction, as illustrated in figure 5.14. This means that the extracted in-plane stresses will be attributed to the area covered by the two transistors. The transistors with current flow labeled $I_1$and I_2 correspond to eq. 3.10 and 3.11 repeated below

$$\frac{\Delta I_1}{I_1} = \pi_l \sigma_l + \pi_t \sigma_t + \pi_v \sigma_v \tag{5.5}$$

$$\frac{\Delta I_2}{I_2} = \pi_t \sigma_l + \pi_l \sigma_t + \pi_v \sigma_v \tag{5.6}$$

where σ_l and σ_t are the longitudinal and transverse in-plane stresses, corresponding in this case to the [110] and [-110] directions, respectively, σ_v is the vertical or out-of-plane stress in [001] direction and π_l, π_t and π_v the corresponding in-plane and out-of-plane piezocoefficients. As there are two equations and three unknowns, being the three normal stresses, this transistor set can extract in-plane stress components only under the assumption that the out-of-plane stress in its surrounding is negligible. The system of equations then reduces to a system with two equations and two unknowns, indicated below

$$\frac{\Delta I_1}{I_1} = \pi_l \sigma_l + \pi_t \sigma_t \tag{5.7}$$

$$\frac{\Delta I_2}{I_2} = \pi_t \sigma_l + \pi_l \sigma_t \tag{5.8}$$

Both transistors in the configuration illustrated in figure 5.14 are undoubtably impacted by out-of-plane stress. The higher the out-of-plane stress, the higher the error introduced when neglecting the out-of-plane stress during in-plane stress extraction using eq. 5.7 and 5.8. Figure 5.15 provides a visual presentation of the share of out-of-plane components in the overall current shift. Extracted in-plane coefficients from the 32nm node long channel channel devices and the only extracted out-of-plane piezocoefficient, from the 65nm node stress test chip are used. The vertical axis in figures 5.15

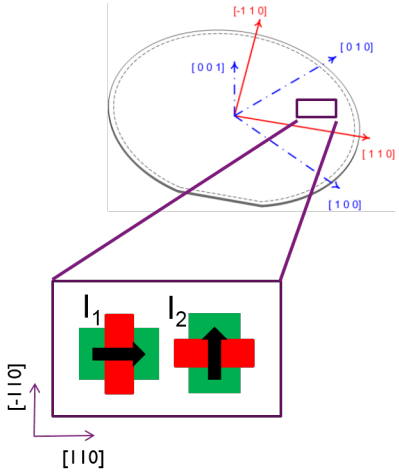


Figure 5.14: In-plane stress in [110] and [-110] direction can be extracted from a two transistor sized area by placing one transistor with current flow in [110] direction and the other next to it with current flow in [-110] direction

a) and b) plots the share of the out-of-plane stress component in the overall current shift as

$$Z = \frac{\pi_v \sigma_v}{(\pi_l + \pi_t)\sigma_i + \pi_v \sigma_v} \cdot 100 \tag{5.9}$$

where the numerator holds the out-of-plane component and the denominator the total contributions of in-plane and out-of-plane stress components. For this case, the two in-plane stresses are taken as equal, $\sigma_l = \sigma_t$, and labeled σ_i. The vertical axis, Z, is plotted in percentage versus in-plane stress, σ_{lt}, and out-of-plane stress, σ_v. Figure 5.15 a) presents the impact of the out-of-plane component for n-type MOSFETs and figure 5.15 b) for p-type MOSFETs.

For a particular value of in-plane stress, the rising out-of-plane stress increases its impact in the equation. As the absolute out-of-plane piezocoefficient value for n-type transistors is similar to the sum of the n-type in-plane piezocoeffients, $|\pi_v| \approx |\pi_l + \pi_t|$, when the out-of-plane stress value reaches the value of in-plane stresses, the out-of-plane stress component $\pi_v \sigma_v$ already accounts for slightly more than 50% of the total current shift. For p-type devices, the Smith out-of-plane piezocoefficient is an order of magnitude lower than the calibrated in-plane piezocoefficients, but the in-plane piezocoefficients are opposite signed, which in case of equal out-of-plane and in-plane stress, leads to a share of the out-of-plane component in the current shift of about 25%, smaller than with the n-type devices.

P-type transistors are somewhat less sensitive to out-of-plane stress, however both transistor types display non-negligible sensitivity to out-of-plane stress. Figure 5.16 plots the out-of-plane stress values which are allowed in presence of in-plane stress if a 10% out-of-plane component share to current shift criteria is taken. For n-type

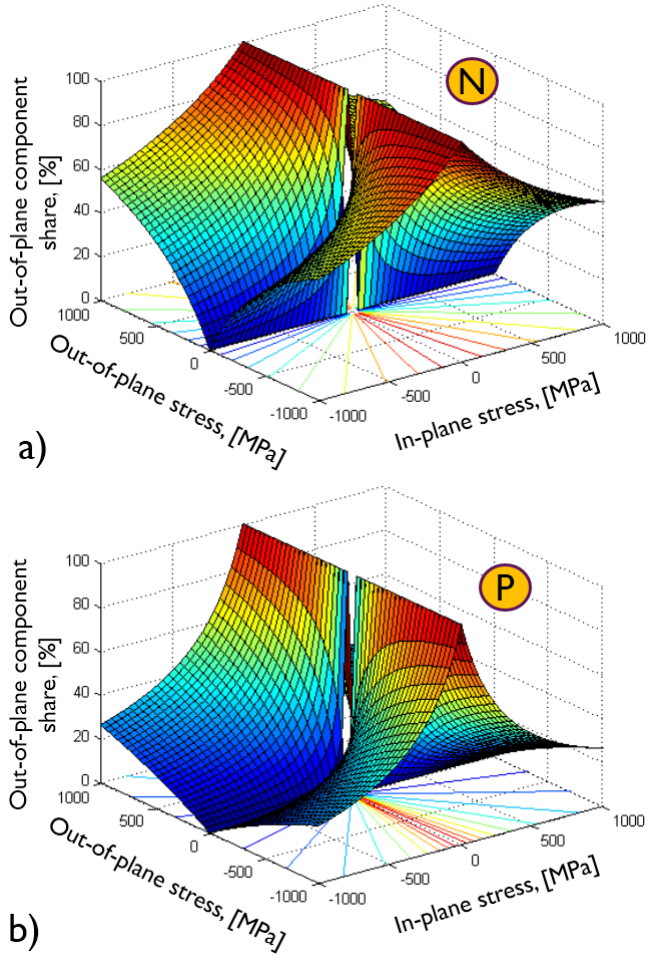


Figure 5.15: Share of out-of-plane component in current shift versus in-plane stress and out-of-plane stress, according to eq. 5.9 for a) n-type devices and b) p-type devices

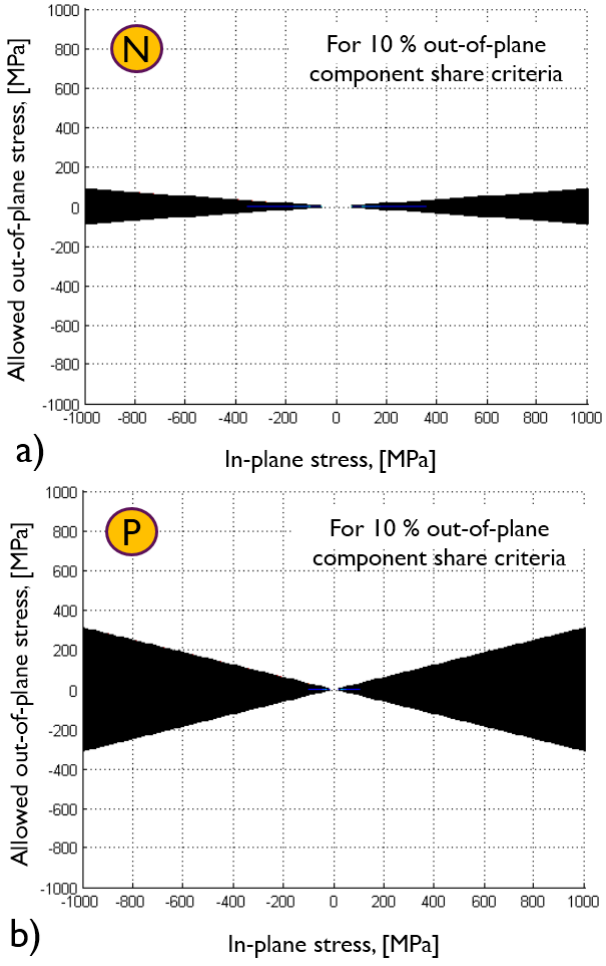


Figure 5.16: Allowed out-of-plane stress values in presence of in-plane stress if a 10% out-of-plane component share in current shift criteria is taken for a) n-type devices and b) p-type devices

transistors, figure 5.16 a), for 10% out-of-plane component share, less than 100 MPa is allowed for 1GPa of in-plane stress. In general for n-type devices, the out-of-plane stress component value should be slightly lower than 1/10 of the in-plane stress value. For p-type transistors, figure 5.16 b), for 1GPa of in-plane stress, a maximum of 300 MPa of out-of-plane stress is allowed for a 10% out-of-plane impact criteria. In general for p-type devices, the out-of-plane stress component value should be slightly lower than 1/3 of the in-plane stress value.

Although limited with vertical stress values, the configuration in figure 5.14 provides a practical way to extract individual in-plane stress components and was used in prototype stacks, results of which will be presented in chapters 7 and 8.

5.3 Assessment of stress impact on advanced devices

5.3.1 FinFETs

Apart from planar MOSFET transistors, the 4-point bending calibration method was applied on post-MOSFET generation devices, 32nm technology node FinFETs, to benchmark their sensitivity to in-plane stress. Figure 5.17 a) presents the layout of one individual tested FinFET. One device consists of 5 fingers, each 20nm wide and 900nm long. The effective fin width per finger is 100nm. Over 5 fingers this equals to a device of a channel width to length ratio of 500/900nm. Figure 5.17 b) gives an SEM image of the FinFET with removed oxide layers. A poly Si gate is visible covering vertically standing Si fins. Figure 5.17 c) shows a cross section of the FinFET showing 3 of its fingers. The vertical Si fins are surrounded by SiO_2 up to the top of the fin where the channel is formed surrounded by metal and poly Si acting as the gate contact. The channel of one fin is magnified in figure 5.17 d). The width of each fin at its top is 20nm. The channel is formed on all 3 sides of the fin where the metal gate and poly Si is present, therefore including also two 40nm side walls, resulting in an effective channel width per finger of 100nm. Five of these fingers therefore equal to a total device channel width of 500nm.

Figure 5.18 a) and b) present in-plane stress calibration results of n-type and p-type FinFETs compared to planar transitors from the same node with the same width to length channel ratio of 500/900nm. Table 5.6 summarizes the piezocoefficient values obtained from the FinFETs and includes the piezocoefficients obtained from planar FETs and Smith's piezocoefficients for comparison. Table 5.7 provides the bias voltages, which were the same for both FinFET and planar devices of the same types. Table 5.8 gives the maximum standard deviation of the obtained piezocoefficient values from n-type and p-type FinFETs and previous planar FETs.

N-type FinFETs, figure 5.18 a), exhibit a longitudinal sensitivity similar to planar FETs. The n-type FinFET mean piezocoefficient is slightly higher, however the maximum standard deviation of n-type FinFETs versus n-type planar FETs, table 5.8, is 3 times higher. When in-plane stress is applied in the transverse direction to the [110] current, the n-type FinFET exhibits approximately 3 times lower sensitivity than its planar counterpart. P-type FinFETs, figure 5.18 b), exhibit lower sensitivity than their planar counterparts in both longitudinal and transverse direction. The p-type FinFET

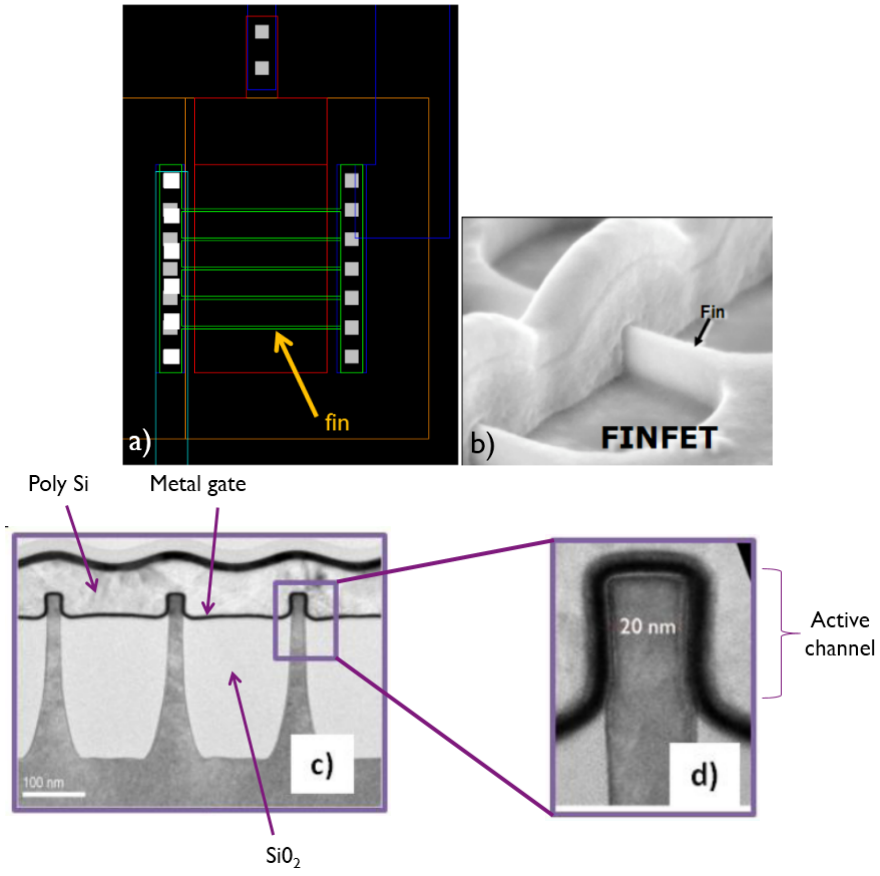


Figure 5.17: In-plane stress calibration of FinFETs: a) layout of one individual 5-finger FinFET, b) SEM image of an individual FinFET with removed oxide layers, two fins covered with poly Si gate are visible, c) FinFET cross section with 3 fingers visible and d) magnification of the top of one fin indicating the FinFET channel

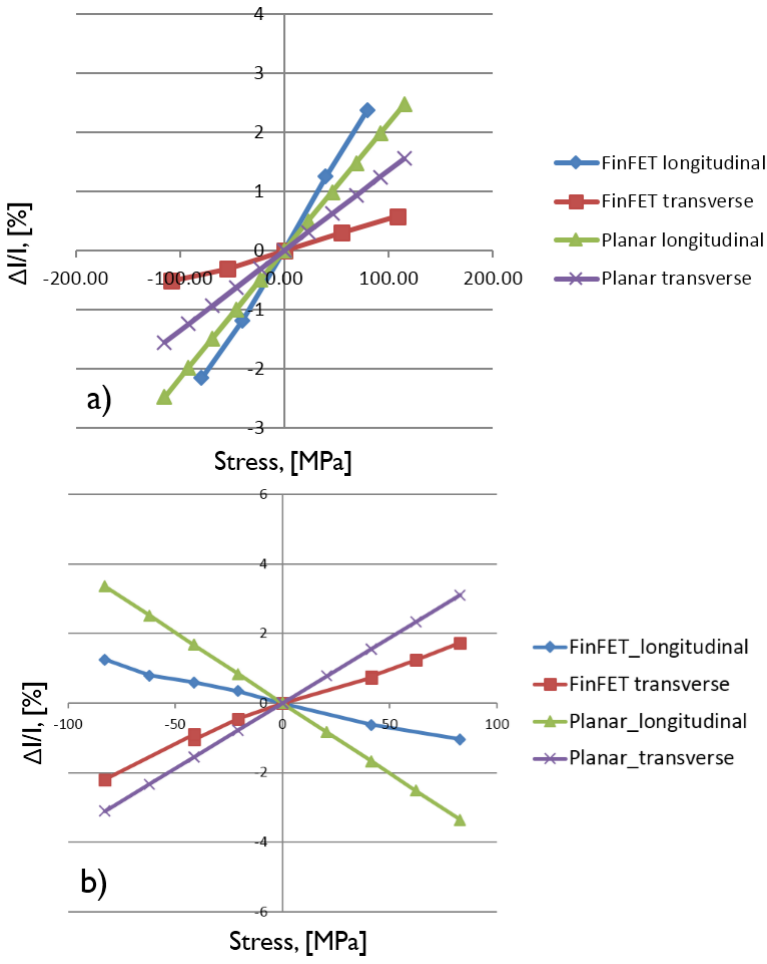


Figure 5.18: FinFET in-plane stress calibration results with included MOSFET calibration results: a) n-type FinFETs and MOSFETs and b) p-type FinFETs and MOSFETs

Technology node	FET type	Width/Length (nm/nm)	Piezocoeff. type	Piezocoeff. Saturation (ppm/MPa)	Smith piezocoeff. (ppm/MPa)
32 nm (FUJI)	Planar, N	500/900	longitudinal	-215	-312
	FinFET, N	500/900	longitudinal	-269	-312
	Planar, N	500/900	transverse	-135	-176
	FinFET, N	500/900	transverse	-42	-176
	Planar, P	500/900	longitudinal	406	718
	FinFET, P	500/900	longitudinal	136	718
	Planar, P	500/900	transverse	-375	-690
	FinFET, P	500/900	transverse	-229	-690

Table 5.6: Summarized in-plane piezocoefficient values of FinFETs with included MOSFET piezocoefficients for comparison

Technology node	FET type	Operating regime	Source (V)	Drain (V)	Gate (V)	Bulk (V)	Power supply (V)
32 nm (FUJI)	N	Saturation	0	1.2	1.2	0	-
	P	Saturation	1.2	0	0	1.2	-

Table 5.7: Bias conditions for FinFETs calibrated to in-plane stress from figure 5.18

Technology node	FET type	Max. standard deviation [MPa]
32 nm (FUJI)	Planar N	$\pm$12
	FinFET N	$\pm$7
	Planar P	$\pm$23
	FinFET P	$\pm$29

Table 5.8: Maximum standard deviation of the obtained piezocoefficients observed on particular transistor types for long channel FinFETs and MOSFETs of the 32 nm technology node

Material	Young's modulus [GPa]	Possion's ratio
Si (orthotropic)	E_{11}= 169 E_{22}= 169 E_{33}= 130	n_{12}=0.064 n_{23}=0.36 n_{31}=0.28
Metal gate	300	0.25
Poly Si	150	0.22
Oxide	50	0.17

Table 5.9: Material properties used in the finite element model from figure 5.19

longitudinal piezocoeffient is approximately 3 times lower than the planar one, and the p-type FinFET transverse piezocoefficient approximately 1.5 times lower. The maximum standard deviation for p-type FinFETs was 4 times higher than for same channel width to length ration planar transistors.

The n-type FinFET exhibited very low sensitivity of 42 ppm/MPa in [-110] transverse direction, with current flowing in [110] direction. The main differences between the channels of a planar MOSFET and FinFET in terms of stress impact on their channels are:

- the channel region in a FinFET is surrounded by metal and poly Si from 3 sides, in a planar FET, the gate materials are present only on one side
- the FinFET current runs not only on the Si surface but through the sidewalls as well

Two possible reasons why the n-type FinFET shows smaller current shift in transverse direction are proposed:

- stress is not fully transferred to the Si region surrounded by gate materials
- the side walls with different Si crystal orientation impact overall sensitivity

In order to investigate stress transfer to the FinFET channel region, a finite element model was built of a single device, presented in figure 5.19., based on the layout from figure 5.17 a). The material properties used in the model are summarized in table 5.9. The FinFET stress distribution was investigated after the device was strained in longitudinal and transverse directions to levels inducing approximately 100 MPa in bulk Si. The simulation was purely thermo-mechanical, with no induced current. Figure 5.20 presents the results of the simulation for both cases of longitudinal and transverse applied stress to the channel region.

In the longitudinal case, when stress is applied in the direction of the fins, stress is distributed equally over the fins including the top channel region. In the transverse case, when stress is applied perpendicular to the fin length, the transfer of stress to the fin is limited. Most of the fin exhibits low stress, however, the channel region still exhibits similar stress as in the longitudinal case. The SiO_2 material surrounding most of the fin is softer than the surrounding materials and could be responsible for

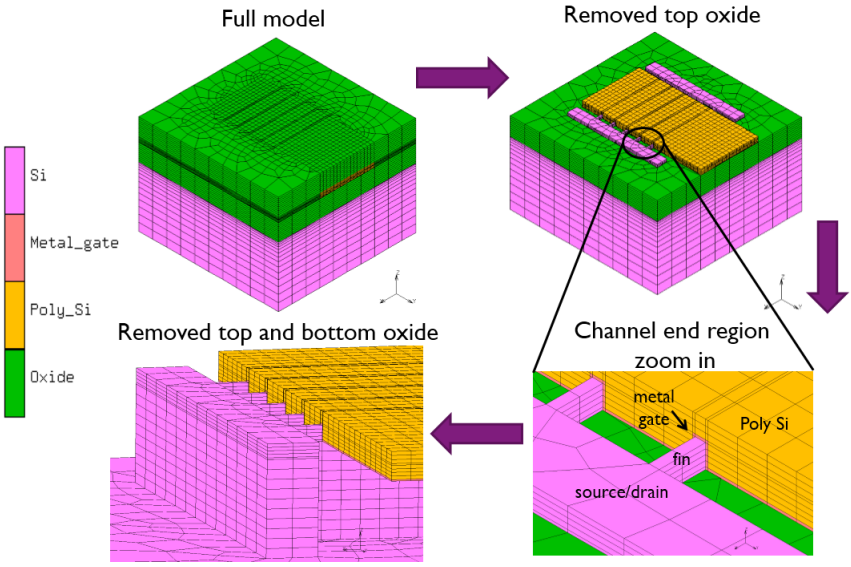


Figure 5.19: Finite element model of a single FinFET from figure 5.17, built to investigate transfer of stress to FinFET channels

low stress transfer to the fin in transverse direction, acting as a stress buffer layer. The poly Si and metal gate are however stiffer materials and transfer stress to the Si channel region they surround. Since only the channel region is responsible for the current flow in a FinFET, according to the finite element model, it is unlikely that the low transverse sensitivity to stress comes from lack of stress transfer to the channel region. It is more likely that the side walls which account for 80% of the channel width play a significant role in the different stress response compared to the planar FET. The [-110] and [1-10] Si surfaces of the FinFET sidewalls have a fundamentally different reaction to stress than the current running through the [001] surface of the top part of the channel. A study involving investigation of stress responses of FinFET sidewalls was not pursued within his PhD, but is being further considered by device technologists. Furthermore, as FinFETs exhibit lower sensitivity than their planar counterparts of the same technology node, they do not show potential for exploitation as stress sensors.

5.3.2 Pseudo-Hall stress sensors

In the last 10 years the pseudo-Hall transistor has been investigated and applied to IC stress measurements [90-92]. It's main claimed characteristics of interest to this work are extraction of shear in-plane stress and the difference between normal in-plane stresses in Si, independent of temperature. The pseudo-Hall operates as a standard MOSFET with the difference that stress is monitored through the stress modified voltage present perpendicular to current flow. This is clarified in figure 5.21.

Figure 5.21 a) and b) display the two measurement types conducted on the pseudo-

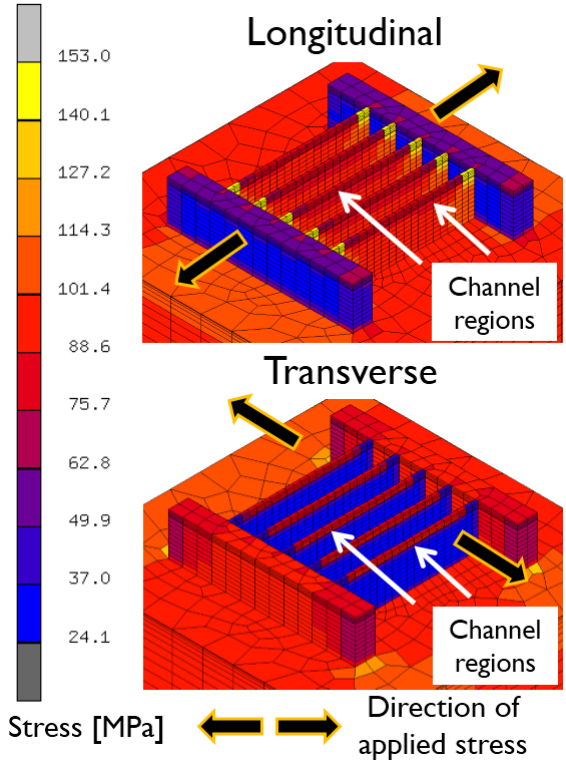


Figure 5.20: Results of the finite element model after application of external strain for approximately 100 MPa in Si, in the longitudinal and transverse direction to the fin. Stress is equally transferred to the channel region in both directions while the fin side walls, irrelevant to current flow, exhibit low stress transfer in transverse direction.

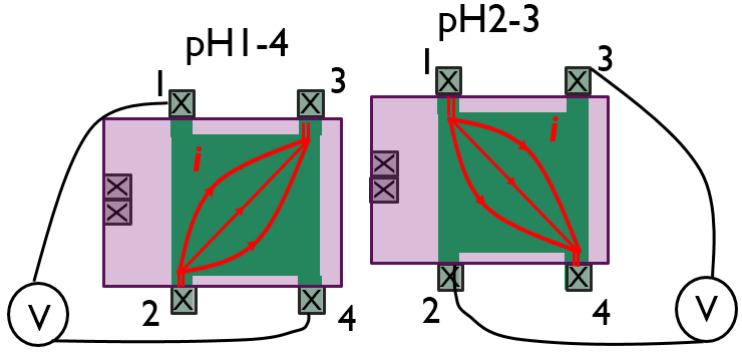


Figure 5.21: Two measurement types performed on the pseudo Hall transistors, swapping the contacts used for source-drain current and pseudo Hall voltage

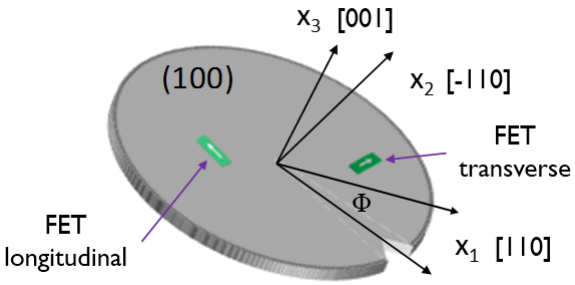


Figure 5.22: Illustration of longitudinal and transverse positions of the transistors on the wafer with respect to the reference frame related to eq. 5.10

Hall transistors. The square shaped active Si area is completely covered with the gate. Four corner contacts can be arbitrarily used to bias the transistor. In the first measurement type, the drain to source current flows diagonally between contacts 2 and 3. The transverse pseudo Hall voltage is measured between contacts 1 and 4. In the second measurement type, the connections are reversed. Drain to source current flows through contacts 1 and 4 while the pseudo Hall voltage is measured on contacts 2 and 3. The pseudo-Hall voltage, V_{pH}, is given as [90-92]

$$V_{pH} = [(\pi_{11} - \pi_{12})\sigma'_{xy} cos(2\phi) - \frac{1}{2}\pi_{44}(\sigma'_{xx} - \sigma'_{yy}) sin(2\phi)]V_b G \qquad (5.10)$$

where π_{11}, π_{12} and π_{44} are the piezocoefficients corresponding to standard Si crystal orientations [100]/[010]/[001] and stress components σ'_{xy}, σ'_{xx} and σ'_{xy} the shear and normal in-plane stress components corresponding to standard device orientation [110]/[-110]/[001], rotated 45° from the standard Si crystal orientations. The drain to source bias is represented through V_b and ϕ is the angle between the device current flow and [110] direction. Figure 5.22 illustrates longitudinal and transverse positions of the transistors on the wafer and the reference frame related to eq. 5.10.

Pseudo Hall transistors were fabricated on the 65nm technology node stress test chip PTCQ. The device reference frame on the PTCQ test chip was in-line with the standard Si crystal orientation, being [100]/[-110]/[001]. Their channel width to length ratio, for both n-type and p-type transistors was 3µm/3µm. N-type transistors were processed rotated 45° to standard Si crystal orientation while p-type transistors were processed parallel to standard Si crystal orientations. This means that the current for n-type transistors was flowing in [100] and [010] directions, depending on the two types of measurements from 5.18 and the current for p-type was flowing in [110] and [-110] directions. According to eq. 5.10, in this configuration, p-type pseudo Hall transistors are most sensitive to normal in-plane stresss and n-type to shear in-plane stress. Figures 5.23 and 5.24 present the stress calibration results for p-type and n-type pseudo Hall transistors, in saturation and linear regime, for both measurement types from figure 5.21. Table 5.10 summarizes the calibration results and table 5.11 the voltage bias used.

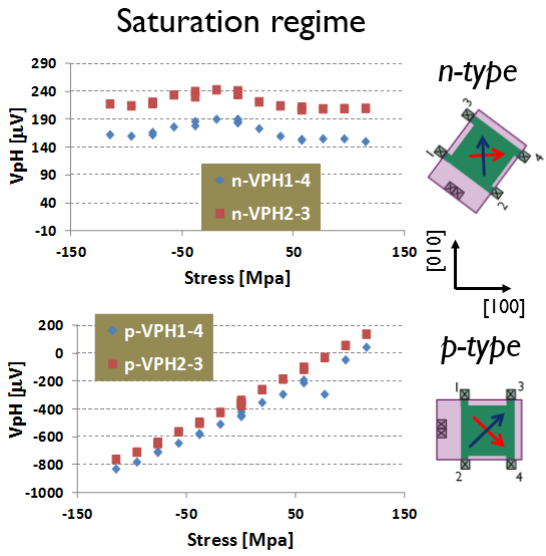


Figure 5.23: In-plane stress calibration results for p-type and n-type pseudo Hall transistors, in saturation regime, for both measurement types indicated in figure 5.21

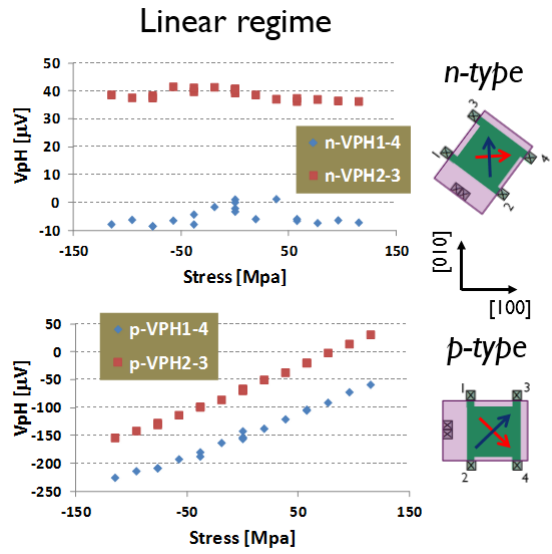


Figure 5.24: In-plane stress calibration results for p-type and n-type pseudo Hall transistors, in linear regime, for both measurement types indicated in figure 5.21

Technology node	FET type	Width/Length (nm/nm)	Piezocoeff. type	Piezocoeff. Saturation (μV/MPa)	Piezocoeff. Linear (μV/MPa)
65 nm (PTCQ)	Pseudo, N	3000/3000	longitudinal	-	-
	Pseudo, P	3000/3000	longitudinal	4.07	0.78

Table 5.10: Summarized in-plane piezocoefficient values of FinFETs with included MOSFET piezocoefficients for comparison

Technology node	FET type	Operating regime	Source (V)	Drain (V)	Gate (V)	Bulk (V)	Power supply (V)
65 nm (PTCQ)	Pseudo N	Saturation	0.05	0.95	1	0	1
	Pseudo N	Linear	0.05	0.1	1	0	1
	Pseudo P	Saturation	0.95	0.05	0	1	1
	Pseudo P	Linear	0.95	0.9	0	1	1

Table 5.11: Bias conditions for pseudo Hall transistors calibrated to in-plane stress from figure 5.23 and figure 5.24

Technology node	FET type	Operating regime	Piezocoeff. type	Piezocoeff. Saturation (μV/MPa)	Piezocoeff. temp. shift (μV/°C)	V_{pH} Temp. shift (μV/°C)
65 nm (PTCQ)	Pseudo P	Saturation	longitudinal	4.07	0.002	0.25
		Linear	longitudinal	0.78	0.0007	0.31

Table 5.12: Summarized in-plane piezocoefficient and absolute V_{pH} sensitivity to temperature. V_{pH} changes with temperature but the stress sensitivity shift with rising temperature is 3 orders of magnitude lower than its response to stress.

In accordance with eq. 5.10, n-type pseudo Hall transistors do not exhibit sensitivity to normal in-plane stress for current flow in [100] and [010] direction for which the angle $\phi = 0$. P-type Pseudo Halls exhibit a sensitivity to stress of 0.78 μV/MPa in linear regime and 4.07 μV/MPa in saturation regime. Their stress sensitivity was assessed at higher temperatures to verify the claimed invariance to temperature.

The transistor was heated up from room temperature to 40ºC and 60ºC through locally applied heat via an air gun followed by in-plane stress application via the 4-pt bending setup, as indicated in figure 5.30. The temperature was monitored on-chip through previously temperature calibrated diodes. The results are summarized in table 5.12. The pseudo Hall piezocoefficient exhibited very low sensitivity to temperature, three orders of magnitude lower than its sensitivity to stress. Although the stress sensitivity seems invariant to temperature, the absolute value of V_{pH} still changes with temperature, in the same order of magnitude as its sensitivity to stress. This means if we wish to be monitor stresses at higher temperatures, the V_{pH} shift due to temperature should be deducted from the total V_{pH} to explore impacts of stress.

5.4 Biaxial stress test

The sensor validation test presented in section 3.2.1, which involved inducing stress by cooling of a die-to-substrate structure, can in fact be used as a biaxial stress test for stress sensors. This test can be used to prove the validity of uniaxially obtained piezocoefficient values in a biaxial stress environment. Figure 5.25 presents the die-to-substrate structure with a die of approximately 10mm x 10mm glued to a 30mm x 30mm plastic substrate. The die edges were aligned with the substrate edges. At room temperature, viscous die attach, STYCAST 50500D, was placed on the plastic substrate followed by picking and placing of the die above. The die attach required 2h of baking at 150 ºC after which the structure was cooled to room temperature causing biaxial stress on the top Si active surface.

Figure 5.26 presents the profile of the die surface obtained with an optical profilometer. The curvature of the die is regular in all directions with a displayed warpage of approximately 40 μm from die center to all four die corners. A stress pattern of the top Si surface obtained with finite element modeling is shown in figure 5.27. The longitudinal and transverse normal in-plane stresses are constant around the middle of the die fading close to the die edges. The position of a sensor array consisting of 130nm technology 800nm/600nm channel width to length ratio nFETs is marked with a red and white rectangle.

To vary the induced stress, in total three die thicknesses were used:

- 725 μm
- 500 μm
- 300 μm

The currents of in total 256 transistors from the array were measured before and after bonding to the plastic substrate. After bonding, at room temperature the devices exhib-

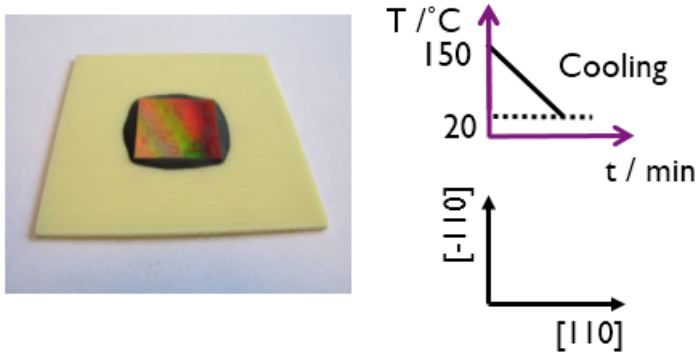


Figure 5.25: The die-to-substrate structure with a die of approximately 10mmx10mm glued to a 30mmx30mm plastic substrate, cooled down from die attach hardening temperature of 150 ºC to room temperature.

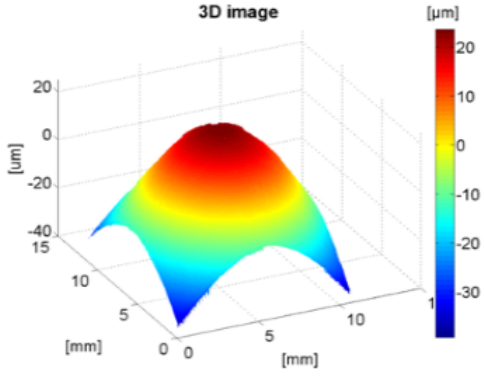


Figure 5.26: Profile of the die surface at room temperature, obtained with an optical profilometer

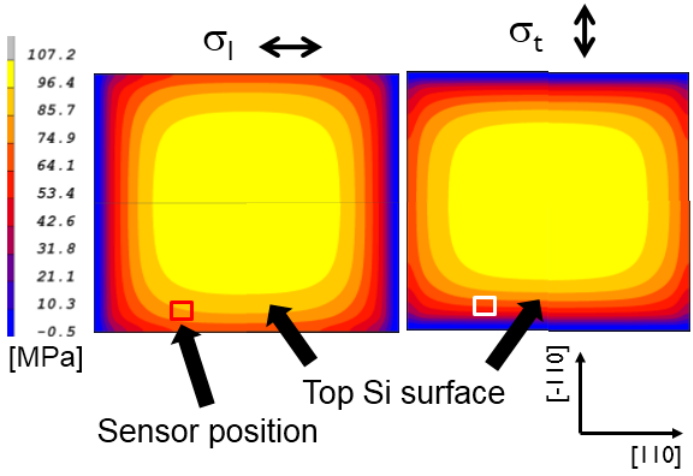


Figure 5.27: In-plane stress pattern of the die top surface, obtained with finite element modeling

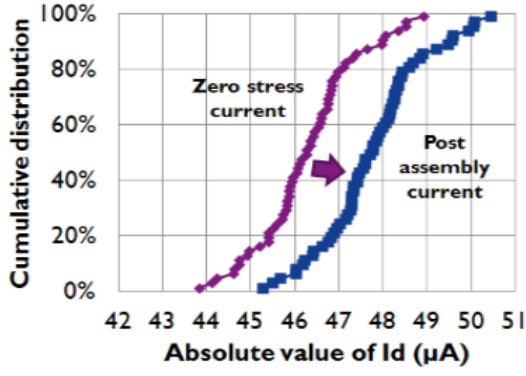


Figure 5.28: Die sensors exhibited a $2\mu A$ shift at room temperature

Die thickness [um]	Die sensors [MPa]	FEM simulation [MPa]	Profilometry [MPa]
725	71.4	73.5	75.5
500	52.5	51	62.5
300	32.9	28.3	32.8

Table 5.13: Extracted stress values from die sensors, FEM and optical profilometry

ited a $2\mu A$ shift of I_D, presented in figure 5.28. Stress was extracted in Si at the sensor position in 3 ways:

- electrical measurements on stress sensors
- finite element model
- by calculating stress from die curvature via Stoney's equation discussed in section 4.4.4

The extracted stress values are summarized in table 5.13 showing good agreement of values from all 3 used methods. This shows that uniaxially obtained piezocoefficient values are valid in a biaxial stress environment. In other words, the reaction of the transistor to stress in a biaxial stress state is a superposition of its reactions in uniaxial stress states.

5.5 High temperature calibration

The setup used, presented in figure 5.29, was used for high temperature calibration of the stress senstivity of MOSFETs and pseudo Halls, previously presented in section 5.3.2. An air gun providing heat was brought in vicinity of the Si die strip clamped within the 4-pt calibration setup. Temperature was monitored on-chip through previously temperature calibrated diodes. Test information is provided in figure 5.30. N-type and p-type MOSFETs of the 32 nm technology node were measured up to 80 MPa in both tensile and compressive direction at room temperature, 50ºC and 60ºC.

Figure 5.31 a) presents stress sensitivity results from n-type MOSFETs. A declining slope is visible with rising temperature. The mean values of n-type stress sensitivities at elevated temperatures is presented in figure 5.31 b). A similar trend with a declining stress sensitivity slope with rising temperature is visible on p-type devices, figures 5.32 a) and b). The results are summarized in table 5.14.

The stress sensitivity to temperature of MOSFETs is considerably higher than the ones observed on pseudo Hall transistors, however both exhibit an absolute value shift of their output with temperature, V_{pH} in pseudo Halls and $\frac{\Delta I}{I}$ drain current shift in MOSFETs, regardless of stress presence. Furthermore, individual normal in-plane stress components, although not temperature compensated can be extracted only with MOSFETs.

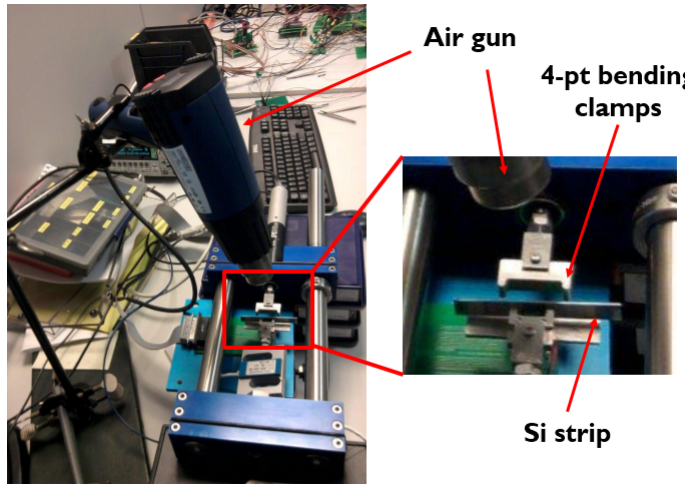


Figure 5.29: Setup used for stress calibration of MOSFETs and pseudo Halls at elevated temperatures involving an adapted 4-pt bending tool and an air gun

Design of experiments	
@Temp [°C]	23, 50, 60
Stress [MPa]	0-80
Devices	N-type and P-type planar FETs
Tech. node [nm]	32
W/L [nm/nm]	500/900

Figure 5.30: Design of experiments for high temperature calibration

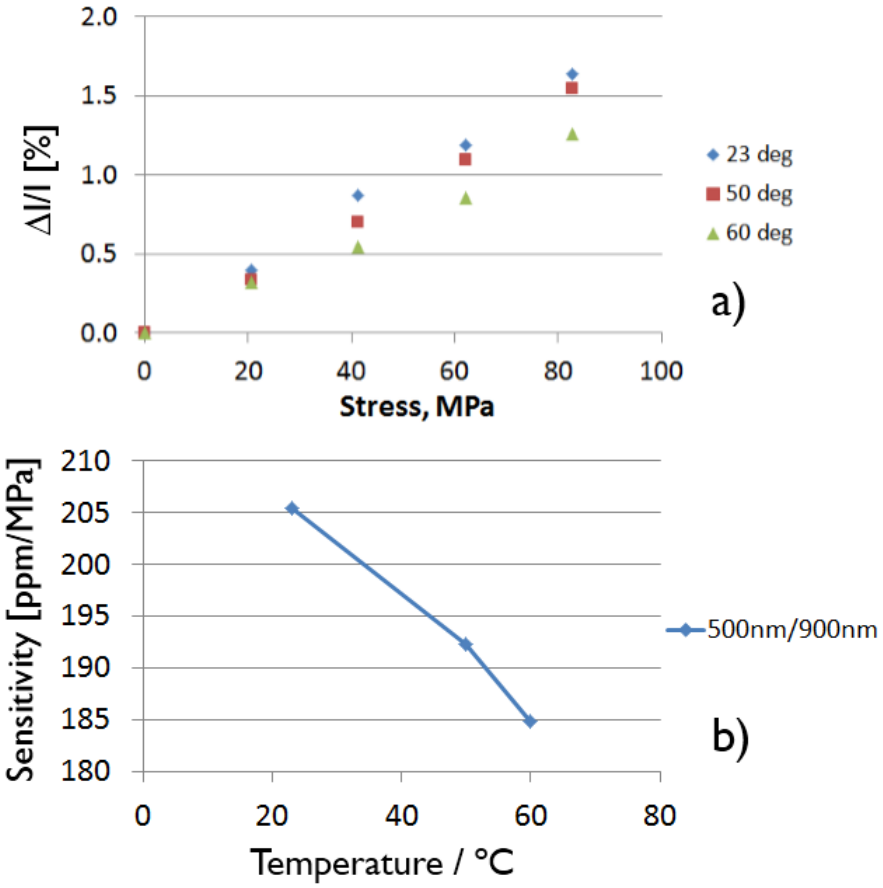


Figure 5.31: Stress calibration at elevated temperatures for n-type MOSFETs: a) changing current shift versus in-plane stress slope and b) declining sensitivity to stress with temperature

Stress sensitivity vs. temperature			
Temp [°C]	23	50	60
n-FET sensitivity [ppm/MPa]	205	192	184
p-FET sensitivity [ppm/MPa]	-397	-311	-259

Table 5.14: Summarized stress sensitivities of 32nm node nFETs and pFETs shifting with temperature

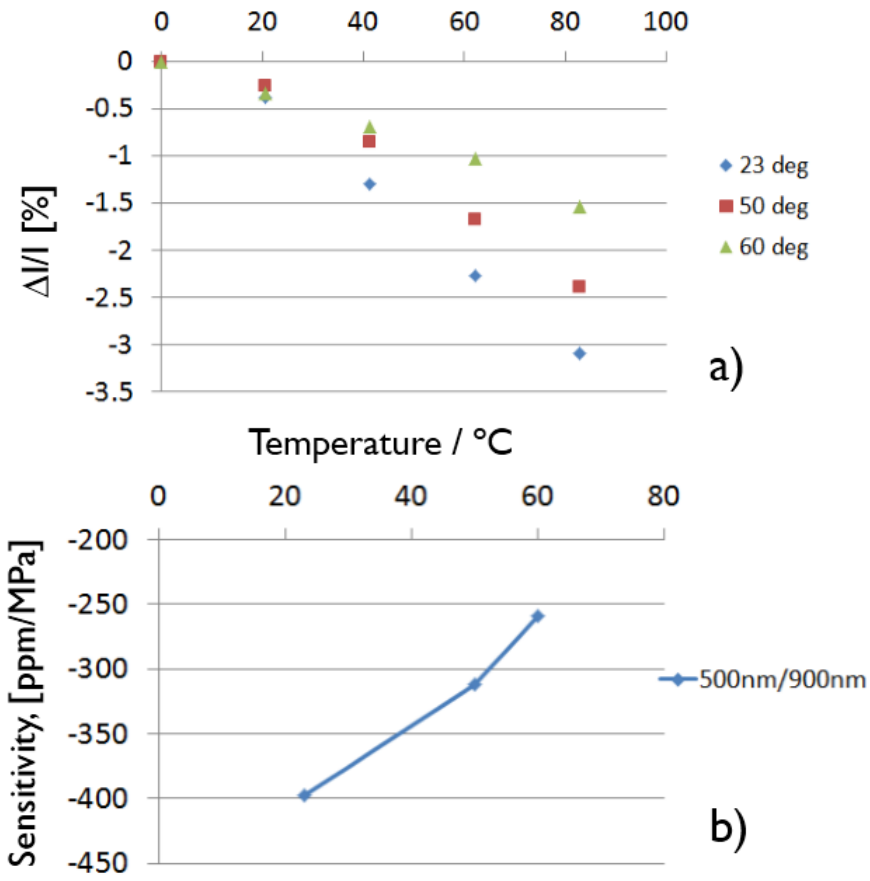


Figure 5.32: Stress calibration at elevated temperatures for p-type MOSFETs: a) changing current shift versus in-plane stress slope and b) declining sensitivity to stress with temperature

5.6 High stress calibration

Stress sensitivity calibration of FEOL devices presented in sections 5.2 to 5.5 was conducted up to maximum 150 MPa. Around 200 MPa of applied in-plane stress, the Si dies often broke. The piezoresistivity model is a linear model and sensitivities to stress of FEOL devices can be expected to behave non-linear at very high stresses, above 1 GPa. In order to reach stress values above 1 GPa, the ultimate tensile strength of Si was assessed. Separation of dies from a wafer via mechanical dicing is known to induce damage to die edges in form of microfractures reducing its ultimate tensile strength (UTS). Mechanical polishing and ion etching was applied in an attempt to increase the Si die UTS.

Small Si strips, 50mm in length, 3.5mm in width with a thickness of 700 μm were saw diced from wafers. Edges of several samples were mechanically polished in attempt to remove damaged layers from the sides. Various polishing paper roughness and polishing time from 10s per side to 1 min per side were applied. Ion etching was applied on one Si side in an etching chamber for 2h using an undisclosed industrial recipe. UTS results were obtained with a 4-pt bending setup. For samples with ion etching stress was applied to the etched Si side. First Si UTS results are summarized in figure 5.33.

All samples that were mechanically polished exhibited lower UTS than the original non-treated samples. One sample that was submitted to ion etching showed an increase in UTS. More samples were made and submitted to ion etching. Figure 5.34 presents ion etching results on 7 new samples. The UTS remains in low ranges if stress is applied to the non-treated side. In total 2 samples returned UTS values similar to non-treated samples and 3 samples exhibited UTS values above the UTS of non-treated samples. One sample was stopped at 1 GPa of applied in-plane stress as the force constraint of the 4-pt bending setup was reached. Although a more thorough study needs to be performed to further assess UTS of Si, ion etching showed promising results which might eventually lead to Si die UTS values above 1 GPa, needed for high stress calibration.

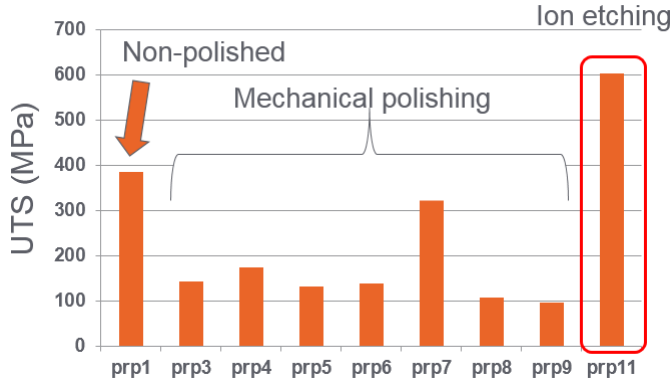


Figure 5.33: Ultimate tensile strength results on Si strips after mechanical polishing and ion etching

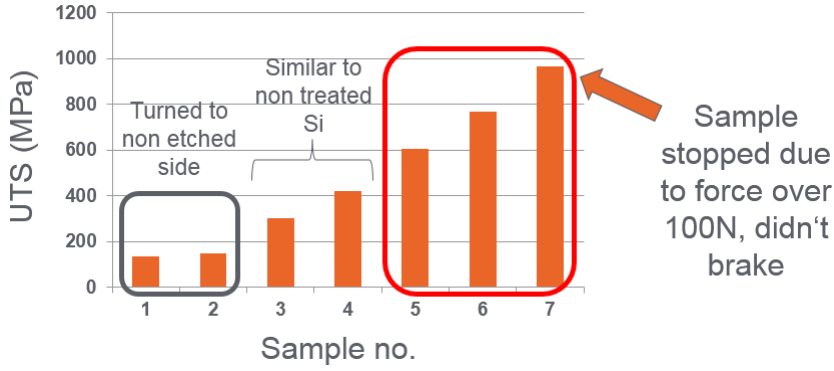


Figure 5.34: Additional ultimate tensile strength results on Si strips after ion etching

5.7 Summary

Stress sensitivity of several FEOL device types and through several FEOL technology nodes was assessed. The stress sensitivities are analyzed from two perspectives:

- To gain information on the nature of stress sensitivity of FEOL devices and its trends through scaling technology nodes
- Their applicability for usage as FEOL CPI stress sensors

MOSFETs were calibrated to in-plane and out-of-plane stress. FinFETs and pseudo-Hall transistors were calibrated to in-plane stress. All calibrations were performed in a [110]/[-110]/[001] Si reference frame. 4-pt bending was used for application of in-plane stress. Out-of-plane stress was applied with a delaminator tip over a Si cube to the Si die and alternatively with the nano-indenter.

The linear piezoresistance model was used to link applied mechanical stress and transistor current shift response. Stress sensitivity is expressed as device piezocoefficients. It is assumed that mechanical stress primarily impacts the mobility of current carriers in transistor channels observable as transistor drain to source current shift. In general, all potential impacts of mechanical stress, apart from mobility shift, that have a linear relation to current shift can be described by the piezoresistance mathematical model. The piezocoefficients are in general device level stress sensitivities, not Si bulk stress sensitivities.

In-plane stress piezocoefficients are obtained from the following devices:

- n-type and p-type MOSFETs, short and long channel, from 3 technology nodes - 130nm, 65nm and 32 nm
- n-type and p-type FinFETs, long channel, from the 32nm technology node
- n-type and p-type pseudo-Hall transistors of one channel size from the 65nm technology node

Out-of-plane stress piezocoefficients are obtained from the following devices:

- n-type MOSFETs, long channel, from the 65nm technology node

MOSFETs, FinFETs and pseudo-Hall devices exhibit a sensitivity to mechanical stress. Smith piezocoefficients for bulk Si are not meant to be used for FEOL devices. Obtained FEOL device piezocoefficients can greatly differ from Smith piezocoefficients for bulk Si. Direct calibration of FEOL devices to stress is necessary.

Summary of in-plane stress calibrations:

- MOSFETs
  - n-type and p-type MOSFETs through all tested technology nodes exhibit sensitivity to stress
  - taking into account their stress sensitivity levels, they are good candidates for usage as CPI stress sensors

  - in the used reference frame, [110]/[-110]/[001], n-type in-plane piezocoefficients are of the same sign and different magnitude, p-type piezocoefficients are of opposite sign and similar magnitude
  - a decrease of sensitivity with newer technology nodes is observed
  - long channel transistors exhibit higher sensitivity than short channel transistors
  - A MOSFET operating in the saturation regime is less senstive to stress than when operating in the linear regime
  - Stress sensitivity of MOSFETs is highly dependent on temperature and drops with temperature increase
  - MOSFETs in a biaxial stress environment reacted equivalently to a superposition of uniaxial stress environments
  - The linear piezocoefficient model might be invalid for high values of stress, approximately above 1 GPa due to non-linear relation between stress and current shift

- FinFETs
  - FinFETs exhibit similar sensitivity in longitudinal direction, but have a distinctively low stress sensitivity in the transverse direction
  - Due to the lower stress sensitivity in transverse direction, they are not recommended for CPI sensors
  - Lower transverse stress sensitivity is most likely not attributed to lower stress transfer in the transverse direction but a fundamental response of FinFET side walls

- Pseudo-Hall transistors
  - Pseudo-Hall transistors present an interesting alternative to transistor current shift monitoring for stress observation, by measuring voltage perpendicular to the drain and source transistor contacts
  - their stress sensitivity is independent of temperature, however the absolute value of the voltage measured to monitor stress still changes with temperature
  - Pseudo-Hall transistors did not exhibit an additional advantage over MOSFETs for implementation as CPI Si stress sensors in 3D IC stacks and packages due to:
    * The stress sensitivity invariance with temperature is not a particular advantage as stress from 3D IC stacks and packages is extracted at room temperature. They could provide more or arguably better performance than MOSFETs for stress extraction at higher temperatures
    * Extraction of individual stress components with pseudo-Hall transistors is not straightforward

Summary of out-of-plane stress calibrations:

- MOSFETs
  - n-type transistors exhibited higher sensitivity to out-of-plane stress in [001] direction than to in-plane stress in [110] and [-110] direction
  - p-type transistors were not calibrated to out-of-plane stress, according to Smith's bulk Si piezocoefficients, p-type Si on [001] Si surface exhibits low sensitivity to out-of-plane stress

An important advantage of MOSFETs is its ability to extract individual stress components. In certain conditions, MOSFETs can be used to extract both in-plane stresses and out-of-plane stress. Shear stress calibration was not attempted on any of the devices. According to piezoresistive theory, Si is not affected by shear stress in [110] and [-110] direction. Mechanical dicing reduces the ultimate tensile strength of Si. Mechanical polishing of Si sample edges did not result in an increase of its ultimate tensile strength. Ion etching of Si samples provided promising results with sample reaching GPa range without breaking. This could enable FEOL device stress calibration at GPa stress levels.

Chapter 6

Back-End-of-Line out-of-plane stress sensors

While chapter 5 included studies on out-of-plane stress sensitivity of planar transistors, in this chapter, an alternative approach for out-of-plane stress extraction is proposed. The possible application of BEOL capacitors as out-of-plane stress sensors is analyzed. The principle of charge-injection-induced error-free (CIEF) charge-based capacitance measurement (CBCM) is presented followed by stress calibration of BEOL capacitors to out-of-plane stress.

6.1 Out-of-plane stress sensors

FEOL devices, planar transistors discussed in section 5.2.2, exhibited sensitivity to out-of-plane stress. Obtaining the stress sensitivity value of FEOL devices can lead to their usage as stress sensors for out-of-plane stress in Si. In this section, extraction of out-of-plane stress in the BEOL has been attempted , by implementing capacitors. The capacitors are simple in form, consisting of two square Cu plates, 30 x 30 μm, one on metal level 1, the other on metal level 2. The general equation for parallel plate capacitors follows

$$C = \varepsilon_0 \varepsilon_r \frac{S}{d} \tag{6.1}$$

where S is the surface of the capacitor plates, d the distance between the plates, ε_0 the permittivity of vacuum and ε_r the relative permittivity of the low-k material between the two Cu plates. In first approximation according to eq. 6.1, application of out-of-plane stress should change the distance between the plates, d, which leads to a capacitance shift of C.

The capacitive Cu plates were processed in the first two metal layers, M1 and M2, of the PTCQ test chip BEOL which consisted of in total 5 metal layers. In order to assure that the capacitance which is being extracted is indeed the capacitance of the sensing plates in M1-M2, the charge-injection-induced error-free (CIEF) charge-based capacitance measurement (CBCM) [119,120] was employed. The capacitance extraction procedure is briefly explained through figure 6.1, 6.2 and equations 6.2, 6.3 and 6.4. The basic electrical circuit, shown in figure 6.1 consists of a p-FET and a n-FET pair in series, on the output of which the capacitive device under test, C_{DUT} is connected. The other contact of C_{DUT} is connected to an external pad. The p-FET and n-FET are controlled by signals V_P and V_N, respectively while the pad connection is controlled through V_{APP}.

The capacitance is extracted in two steps:

- Charging the capacitance under test, C_{DUT} and the parasitic capacitance C_{par}
- Discharging the parasitic capacitance C_{par}

The steps are visualized by the clock signals in figure 6.2. In the first step, the probe pad is grounded. This leads to repeated charging and discharging of both the C_{DUT} and C_{par} between V_{dd} and ground. The charging current through the top p-FET is expressed as

$$I_1 = (C_{DUT} + C_{par}) \cdot V_{dd} \cdot f \tag{6.2}$$

where V_{dd} stands for the power supply connected to the top contact of the p-FET and f is the frequency of the non-overlapping clock signals. In the second step, a controlled pulse is applied to the external pad. In t_1 the pad pulse, V_{APP}, rises up from ground to V_{dd} before the p-FET is turned on. Grounding the signal V_P after t_2, opens the p-FET pulling the output node to V_{dd}. In this phase, only the C_{par} will be charged as

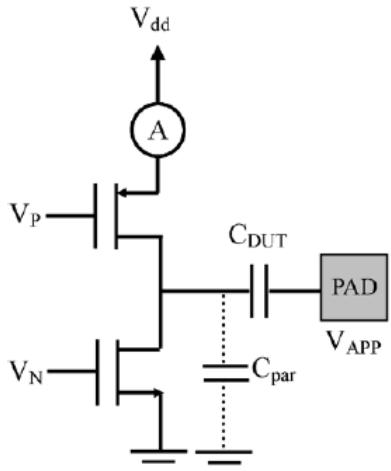


Figure 6.1: The basic electrical circuit of the charge-injection-induced error-free (CIEF) charge-based capacitance measurement (CBCM)

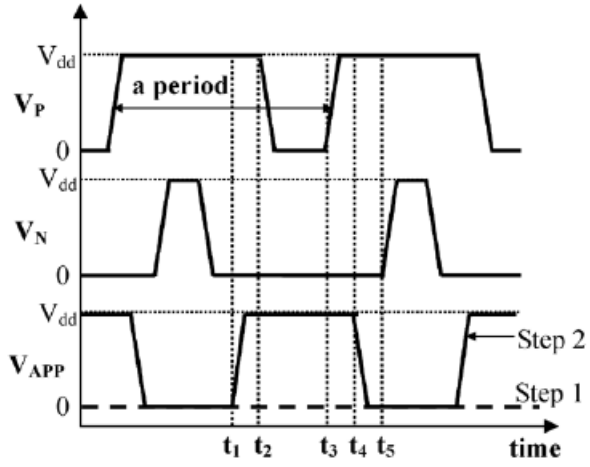


Figure 6.2: Clock signals corresponding to electrical circuit on figure 6.1 visualizing the capacitance extraction procedure

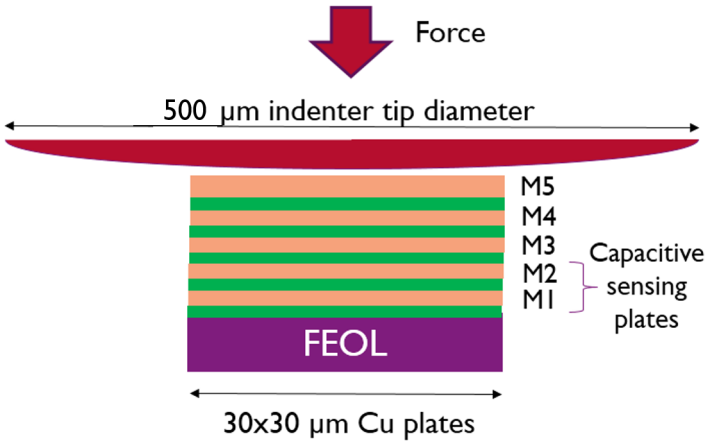


Figure 6.3: Illustration of the out-of-plane calibration of the capacitive sensors with the 500 μm in diameter spherical tip of nano-indenter tool. The capacitive plates are positioned on M1 and M2 levels.

there is no potential drop over C_{DUT}. Before the n-FET is turned on in t_5 to discharge the output node, the pad pulse is grounded in t_4. The current that charges the parasitic capacitance after the V_P drop is measured through the p-FET as

$$I_2 = C_{par} \cdot V_{dd} \cdot f \tag{6.3}$$

The desired capacitance C_{DUT} can then be calculated as

$$C_{DUT} = \frac{I_1 - I_2}{V_{dd} \cdot f} \tag{6.4}$$

The out-of-plane calibration was conducted with the 250 μm in diameter spherical tip of nano-indenter tool, omitting the Si cube which was used for planar transistors, as illustrated in figure 6.3. The nano-indenter tool and spherical tip are described in section 4.3. The sensing capacitor is processed in metal layers $M1$ and $M2$. Figure 6.4 presents measurements of the first phase current, I_1, versus frequency under several loads from the nano-indenter. The linear response at all frequencies acts as a verification of proper device operation, in accordance with eq. 6.2. Figure 6.5 presents the calibration results of the capactive plates. The intial capacitance, C_0, was measured to be 220 fF. Apart from one point, the capacitance shift exhibited a nearly linear response to applied compressive out-of-plane force from the nano-indenter. For force slightly above 600 mN, the capacitance shift rose to just above 8%. The stress sensitivity of the device versus force was approximated to 132 ppm/mN. This value is extracted from tests on a single sample.

As with other stress sensing devices from chapter 5, the sensitivity value should be transferred to a more understandable measure of ppm/MPa. Transferring force to stress is challenging in particular as the force is applied through a spherically shaped

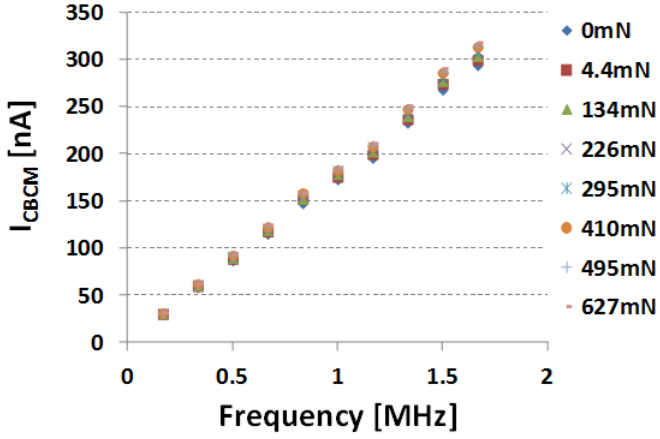


Figure 6.4: Measurements of the first phase current, I_1, versus frequency and under several loads from the nano-indenter. The linear response at all frequencies acts as a verification of proper device operation, in accordance with eq. 6.2.

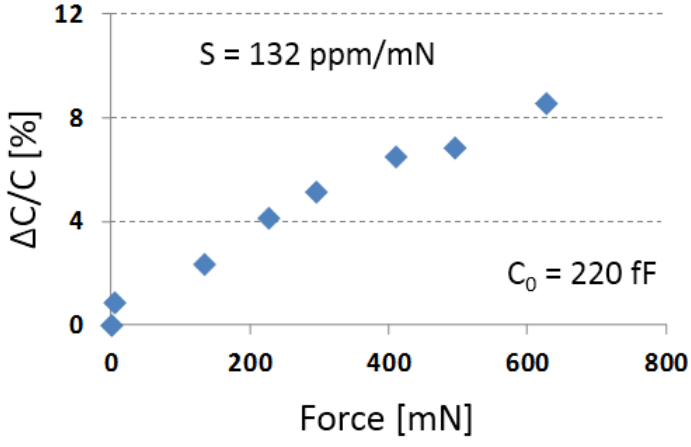


Figure 6.5: Calibration results of the capactive plates. The intial capacitance, C_0, was measured to be 220 fF. The capacitance shift exhibited a linear response to applied compressive out-of-plane force from the nano-indenter.

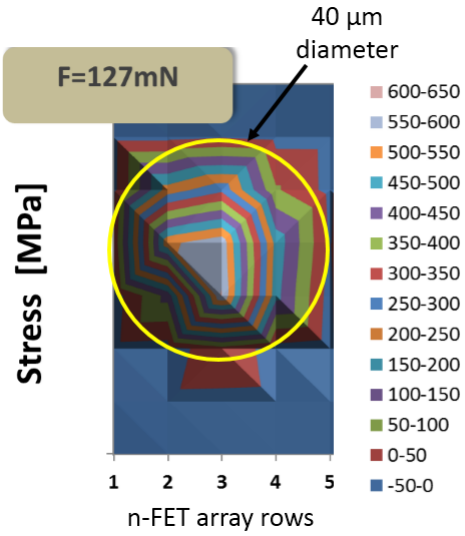


Figure 6.6: The reaction of part of the n-FET array area at an applied out-of-plane force of 127 mN, used to calculate the effective impact area of the nano-indenter spherical tip. Approximately a 5x5 subarray of nFETs from the total 7x7 nFET array was affected. The affected area was approximated to 40 μm in diameter.

tip. The force itself can be read from the nano-indenter tool internal sensors, however the area that the force is applied to is not straightforwardly obvious. Although the 500 μm in diameter spherical tip is almost an order of magnitude larger than the 30x30μm Cu plates, the curvature of the tip cannot be neglected. The exact contact area is hard to obtain with certainty. In an attempt to transfer the applied force to stress, a workaround was attempted to obtain a first approximation of the stress sensitivity in ppm/MPa. The nano-indenter force was applied to the same n-type transistor array from the out-of-plane stress calibration in section 5.2.2, this time without the usage of the Si cube. Figure 6.6 displays the reaction of part of the n-FET array area at an applied force of 127 mN. Approximately a 5x5 subarray of nFETs from the total 7x8 nFET array was affected. The affected area was approximated to 40 μm in diameter. The average stress in the 5x5 nFET subarray equals to 85 MPa. In this simplified approach, this would imply that a force of 127 mN applied with the spherical tip causes an average stress of 85 MPa on an 40 μm in diameter area in the FEOL. In terms of the Cu plates in the BEOL, if assumed that minimal force is transfered to the low-k material, the 30x30 μm Cu plates will take all the out-of-plane stress. Then the 127 mN of applied out-of-plane compressive force will cause a stress of 102 MPa over the 30x30 μm Cu plate area. Having transfered now the nano-indenter force to stress, the stress sensitivity of the BEOL Cu plates now equals to 153 ppm/MPa. Although the transfer from force to stress included several approximations, the final stress sensitivity provides a feeling of the range of sensitivity and an initial value that could be used in practice until a more detailed approach is developed. The force and stress sensitivity values of the BEOL Cu

Out-of-plane sensitivity	Initial calibration [ppm/mN]	Transfer to stress [ppm/MPa]
BEOL Cu plates	132	153

Table 6.1: The final force and stress sensitivity values of the BEOL Cu plates

plates are summarized in table 6.1.

6.2 Summary

Capacitive BEOL plates are proposed as out-of-plane stress sensors as an alternative to FEOL devices. Their capacitance is extracted with the charge-injection-induced error-free (CIEF) charge-based capacitance measurement (CBCM) method. The BEOL plates were calibrated to out-of-plane stress with a nano-indenter using a spherical tip and exhibited a sensitivity of 132 ppm/mN. Calculation of applied stress to the capacitive plates with the nano-indenter using a spherical tip based on known applied force monitored by the nano-indenter is not straightforward. Approximating the effective impact surface of the spherical tip, a sensitivity of the BEOL plates was calculated to be 153 ppm/MPa. The device proves to be sensitive to stress and can be considered as a BEOL stress sensor. Further work needs to be done to assess its applicability and implementation.

Chapter 7

3D stacked ICs

Chapter 7 comprises results from studies on 3D IC stacks. The effects of stress on the Si die FEOL caused by 3D IC stacking are monitored on three different generations of stacks: based on the ETNA, FUJI and PTCQ test chip. Data is obtained primarily with electrical measurements of stress sensors distributed on the Si dies in the 3D IC stacks and correlation with finite element simulations. Section 7.1 introduces the basics of the observed local stress mechanism in 3D IC stacks, the underfill-microbump stress mechanism, through the ETNA stack. The implications of the underfill-microbump interaction mechanism on IC layouts are discussed in section 7.2. Section 7.3 presents a study of mitigation of the underfill-microbump stress. Guidelines for reducing stress generated by the underfill-microbump interaction are given. Section 7.4 analyzes stress from subsequent generations of stacks, the FUJI and PTCQ stack. On the FUJI stack, section 7.4.4, stress in Si is analyzed on the opposite side of the microbump, at room temperature and elevated temperatures. On the PTCQ stack, section 7.4.5, stress in Si is analyzed over the whole die, on the opposite side of the microbump and between microbumps, at room temperature.

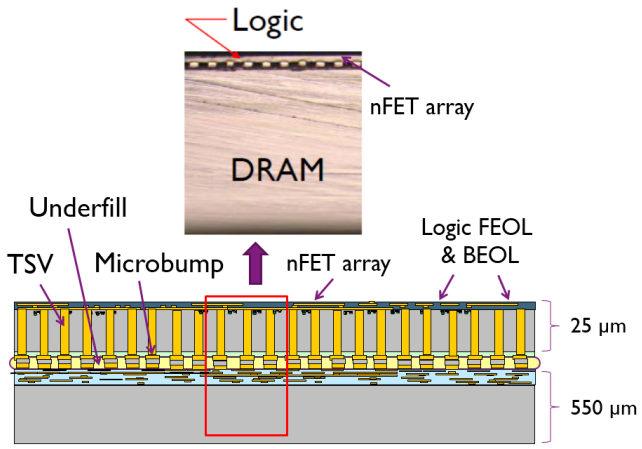


Figure 7.1: Illustration of the first generation 3D IC stack prototype, a functional logic on DRAM structure. The logic die was represented by the ETNA test chip.

7.1 Basics of the underfill-microbump interaction mechanism

The underfill-microbump interaction mechanism is an occurence of mechanical stress in the Si die after 3D IC stacking, in the surrounding of microbumps in the presence of underfill. This chapter focuses on:

- the origins of the stress in the FEOL generated by this mechanism
- the implications of the underfill-microbump interaction mechanism on FEOL devices
- mitigation of underfill-microbump generated stresses

At the beginning of this thesis, no literature was found comprising a study of the origins of stress in the FEOL after 3D IC stacking or 3D IC packaging in connection with underfill and microbump interaction. The focus of the chapter is the stress occuring in Si on the opposite Si side of the microbump.

A 3D IC stack was produced consisting of a functional logic on DRAM structure, presented in figure 7.1. The 10.5 mm x 10.5 mm logic die, labeled ETNA, processed in 130 nm technology was thinned down to 25 µm and penetrated with 5 µm in diameter TSVs with an aspect ratio of 5:1. A 550 µm thick 8.5 mm x 7.5 mm memory die was stacked on the logic die back-to-face by means of thermo-compression bonding at 250 ºC. The gap between dies measured 13 µm consisting of disclosed underfill material and Cu-Sn based microbumps with 5 µm Cu pads on both sides and 3 µm of Cu-Sn intermetallic in between. The microbumps, 30 µm in diameter, were separated at a distance of 250 µm. The image of the 3D stack cross section above the illustration in figure 7.1 emphasizes just how thin the logic die is compared to the memory die in real

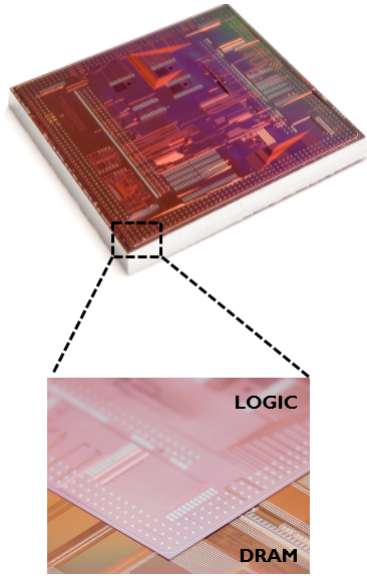


Figure 7.2: Photograph of the assembled logic on DRAM stack from figure 7.1. The DRAM is slightly larger than the logic die.

scale. In further reference, the logic die will be considered as the top die with its active region being on its top side. Figure 7.2 provides a photograph of the assembled logic on DRAM stack, from a case where the DRAM was slightly larger than the logic die.

A 16 x 16 array of n-type FETs with channel width to length ratio 800 nm x 600 nm was positioned above the microbumps in absence of a TSV, in the thin logic die FEOL, as indicated in figure 7.3. The transistor array covered an area of approximately 62 µm x 37 µm. With the microbump of 30 µm in diameter, positioned below the center of the array, the full projected surface of the microbump was covered with n-type FETs plus an additional 16 µm of the array from each lateral side. A reference array was copied at a distance far away from the microbump. The position of the n-type FET arrays can be viewed also in the illustration in figure 5.1.

The currents of the n-type FET array in the thin logic die were measured in saturation before die thinning and after die thinning and stacking, with and without the presence of the underfill. Figure 7.3 a), in red points, presents the measured current shifts from microbump nFET array row 6 of 16 passing close to the center of the array and the underlying microbump, normalized to the values of the reference array. The blue points present the measured current shifts from the microbump nFET array in absence of the underfill. It was revealed that the current of the nFETs above the microbumps in the presence of underfill shifted by more than 40%. Without the underfill, the nFETs exhibit negligible current shift.

A finite element model of the stack is employed in MSC Marc [107] to further investigate this process. Table 7.1 comprises the material properties used. Materialwise,

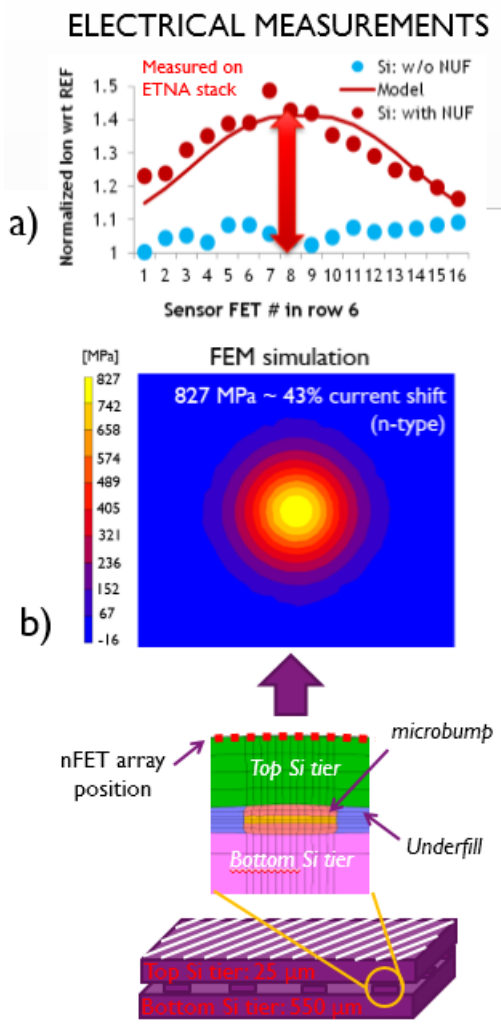


Figure 7.3: The currents of the n-type FETs within the array in the thin logic die were measured in saturation before and after stacking, with and without the presence of the underfill. Figure 7.3 a), in red points, presents the measured current shifts from microbump nFET array row 6 of 16 passing close to the center of the array and the underlying microbump, normalized to the values of the reference array. The blue points present the measured current shifts from the microbump nFET array in absence of the underfill. The full red line present current shifts obtained by finite element modeling of the structure. Figure 7.3 b) presents the major principal stress obtained in Si by finite element modeling of the structure on the top surface of the FEOL logic die, opposite of the microbump. The maximum stress in Si above the microbump was simulated to be 827 MPa, in the center of the circular stress pattern, corresponding to the projected center of the underlying microbump.

Material properties	Si	Cu	Cu-Sn	Underfill
Young's modulus [GPa]	169	117	108	$T<T_g$ 2-8 $T>T_g$ 0.1
CTE [ppm/MPa]	2.6	16	19	$T<T_g$ 30-80 $T>T_g$ 100+

Table 7.1: Material properties in the 3D IC stack environment

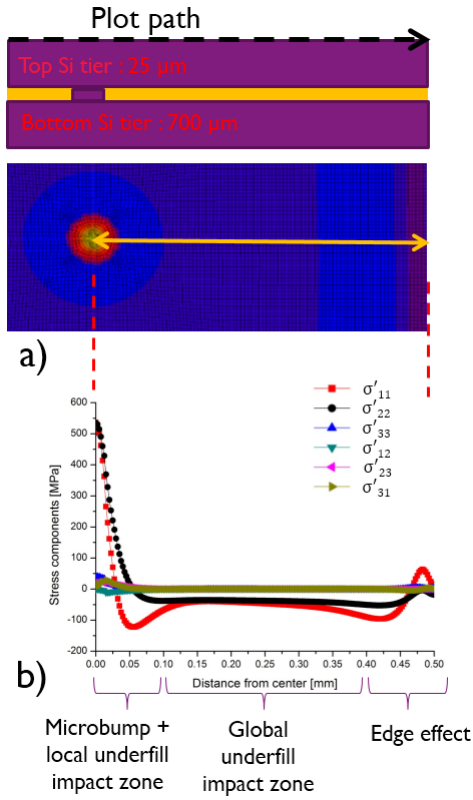


Figure 7.4: Distribution of individual stress components in the Si on the top side of the logic die above the underlying microbump, as in figure 7.3. The in-plane stress components are dominant. Three regions of the underfill-microbump stress mechanism can be distinguished. The circular shape above the graph indicates the major principal stress pattern on the top Si surface.

the model consisted of two Si dies without the BEOL or TSVs, connected with Cu-Sn microbumps and underfill around. All materials were modeled as isotropic and Poisson's ratio for all materials was set to 0.3. Cu, Sn and Si exhibit a Poisson's ratio around 0.3. The Poisson' ratio of the epoxy based underfill was unknown and estimated to 0.3 as well due to the fact that epoxy materials often exhibit Poisson's ratio around that value. The Cu-Sn properties in table 7.1 refers to the intermetallic compound created from the interaction of Cu and Sn. It is considered that the Sn completely transformed into a Cu-Sn intermetallic. Furthermore, Cu is modeled as a plastic material with a piecewise linear function, an idealized work hardening stress/strain curve, as presented in figure 4.4. The yield stress of Cu was set to 172.3 MPa.

Figure 7.3 b) presents the major principal stress obtained in Si from the top surface of the FEOL logic die, opposite of the microbump. The stress pattern in Si occuring from the underfill-microbump interaction is circular with highest stress in the center of the circular shape, fading from the center of the microbump. The maximum stress in Si above the microbump was simulated to be 827 MPa, in the center of the circular stress pattern, corresponding to the projected center of the underlying microbump. The major principal stress equals to the individual normal in-plane stress components, σ_l and σ_t. Eq. 5.7 and 5.8 were implemented within MSC Marc transfering the finite element obtained stresses to current sfhits for n-type of p-type Si. The piezocoefficient values were taken from calibration of particular transistors from chapter 5. In this case, n-type long channel MOSFETs from the 130 nm technology node, ETNA test chip, were utilized. Using the piezocoefficient values from table 5.1 and figure 5.3, the total current shift sums up a value of 43%, matching the electrical measurements. The current shift results of the finite element model corresponding to the position of the 6th row in the nFET array are added to figure 7.3 a), showing good matching.

Figure 7.4 shows the distribution of individual stress components in the Si on the top side of the logic die above the underlying microbump, as in figure 7.3. The two in-plane stress components, $\sigma_{11}^{'}$ and $\sigma_{22}^{'}$, in sensor terms corresponding to σ_l and σ_t respectively, are dominant while the other stress components are negligible. Therefore, in the case of stress in Si on the opposite side of the microbump, it is safe to neglect out-of-plane stress as well. Since MOSFET stress sensors are sensitive to out-of-plane stress, as discussed in section 5.2.2, this goes in favor of in-plane stress extraction, as discussed in section 5.2.3. Furthermore, 3 stress regions can be detected:

- microbump + local underfill impact zone
- global underfill impact zone
- edge effect

The first region denotes the area where the underfill and microbump interact causing main stress in Si. The global underfill impact zone relates to a small amount of compressive stress generated without interaction with microbumps due to interaction of Si and the shrinking underfill itself. The last area is an artefact of the simulation with proximity of the stress to the edge of the model.

Figure 7.5 presents the two dimensional stress patterns of the single microbump in-plane stresses in figure 7.4, above a 5x5 microbump array. Figure 7.5 shows 1/4 representations of the 5x5 microbump array. Each oval shape corresponds to one underlying

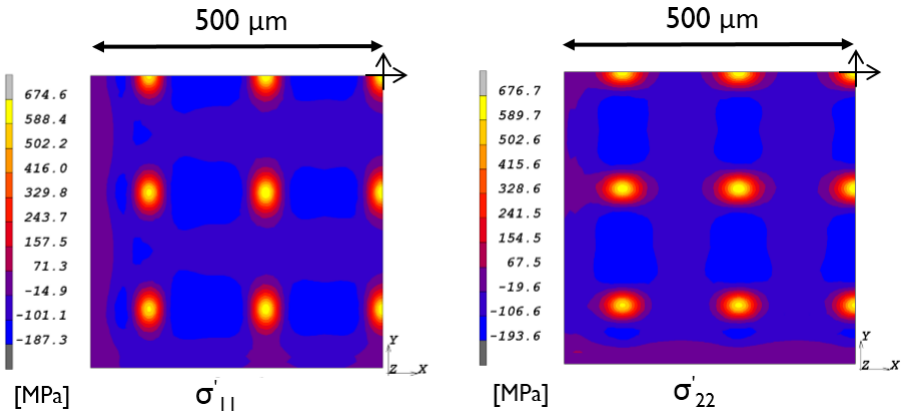


Figure 7.5: Two dimensional representation of the single microbump in-plane stresses from figure 7.4, above an array of microbumps. The images show 1/4 representations of the top Si surface above of a 5x5 microbump array. Each oval shape corresponds to one underlying microbump. These two in-plane stress components sum up to a major principal stress shape as seen in figure 7.3 b).

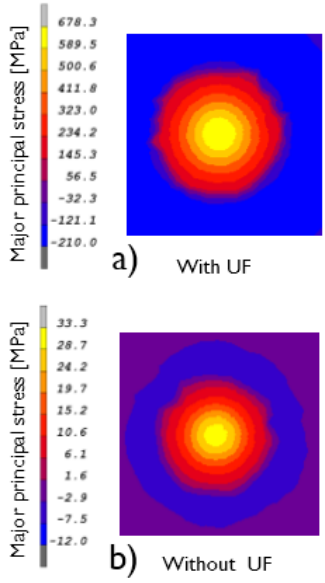


Figure 7.6: Results of the finite element model with and without the presence of underfill. In the absence of underfill, figure 7.6 b), stress is an order of magnitude lower than with included underfill, figure 7.6 a).

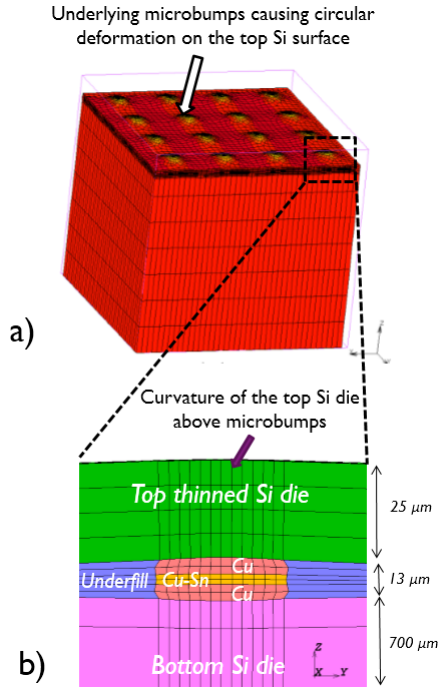


Figure 7.7: Figure 7.7 a) presents a full model of a 3D IC stack with 25 μm thick Si stacked on 700 μm thick Si after cooling from thermocompression bonding temperature of 250 ºC to room temperature, 25 ºC. The exaggarated displacements reveal local warpage of the thin Si die above the underlying microbumps. Figure 7.7 b) zooms in on a cross section of one microbump area. The thin Si is being bent over the microbump

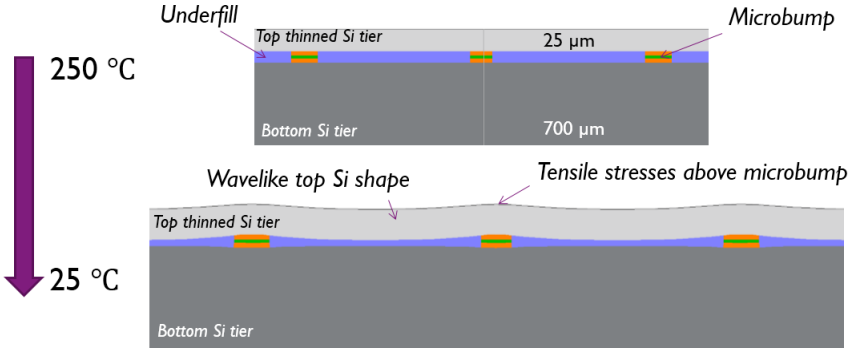


Figure 7.8: Illustration of the underfill-microbump stress generation. The underfill has by far the highest CTE than any material in its surrounding, visible also from table 7.1. During cooling of the 3D stack from bonding temperature of 250 ºC to room temperature, the underfill shrinks more than surrounding materials and pulls the flexible thin Si over the microbumps causing a local curvature in the microbump vicinity

microbump. These two in-plane stress components sum up to a major principal stress shape as seen in figure 7.3 b).

Underfill properties are not specifically mentioned due to a non-disclosure agreement with the material suppliers. Underfill properties were obtained from DMA and TMA curves or alternatively official datasheets when DMA and TMA curves were not available. The absolute values of stress between figure 7.3 b) and figures 7.4 and 7.5 differ as a different underfill material was used causing slightly lower stress. These simulations were performed before knowing the exact underfill which was used in the stack. The impact of underfill properties on stress values will be covered in section 7.3.

Electrical measurements in figure 7.3 a) exhibit negligible current shift in absence of the underfill, compared to the large current shifts in the presence of underfill. Figures 7.6 a) and b) present the results of the finite element model with and without the presence of underfill. In the absence of underfill, figure 7.6 b), stress is an order of magnitude lower than with included underfill, figure 7.6 a).

The origin of the underfill-microbump mechanism creating the stresses shown in figures 7.3 to 7.6 can now be explained in more detail. Figure 7.7 a) presents a full model of a 3D IC stack with 25 μm thick Si stacked on 700 μm thick Si after cooling from thermocompression bonding temperature of 250 ºC to room temperature, 25 ºC. The exaggarated displacements reveal local warpage of the thin Si die above the underlying microbumps. Figure 7.7 b) zooms in on a cross section of one microbump area. The thin Si is being bent over the microbump.

Figure 7.8 reveals the mechanism in the background. The underfill has by far the highest CTE compared to any material in its surrounding, visible also from table 7.1. During cooling of the 3D stack from bonding temperature of 250 ºC to room temperature, the underfill shrinks more than surrounding materials and pulls the flexible thin Si over the microbumps causing a local curvature in the microbump vicinity. The stretched and bent Si now exhibits mechanical stress, tensile in-plane stress on the top Si surface opposite of the microbump. Finite element simulations from figure 5.10, also suggest compressive in-plane stress and compressive out-of-plane stress in Si on the microbump side. Since Si is a piezoresistive material, the mechanical stress reflects into mobility shift of current carriers, electrons and holes resulting in current shift of Si based FEOL devices such as MOSFETs.

7.2 Impact on 3D IC layout

The generated stress in the FEOL from the underfill-microbump interaction mechanism causes a current shift of n-type and p-type FEOL devices. FEOL devices have varying sensitivities to stress, as presented throughout chapter 5, depending on the orientation of the device on the Si wafer and the direction of occuring stress with respect to the direction of the device electrical current. The acting underfill-microbump mechanism stress field in Si has a certain impact zone in which it can affect the current of FEOL devices. In the case of logic on DRAM from section 7.1, the underfill-microbump impact zone, figure 7.4 b) extended to approximately 140 μm from the microbump center. Depending on the obtained stress sensitivity of FEOL devices and the known levels of stress in the vicinity of the microbump, keep-out zones can be established

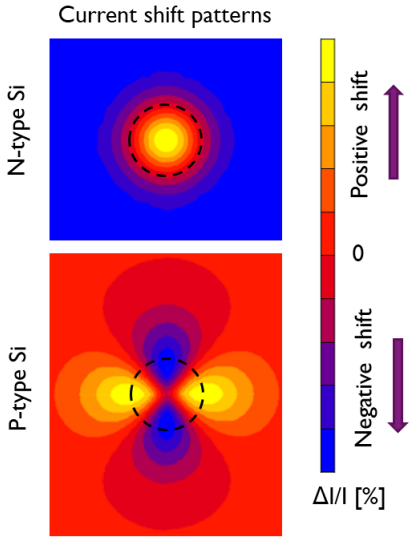


Figure 7.9: The occuring current shift patterns in Si from the underfill-microbump mechanism on the opposite side of the microbump obtained by finite element simulation. The projected contours of the underlying microbump are plotted on the current shift patterns with dashed circular lines. The current shift patterns from figure imply what effect a n-type or p-type transistor will exhibit if placed in these regions.

during IC design development. These keep-out zones can act as proximity rules during IC layout production as regions in which devices should not be placed in order to contain the current shifts of FEOL devices within desired levels. This section deals with current shift patterns in Si and possible keep-out zones for FEOLs positioned on the thinned Si die opposite of the microbump, as in the logic on DRAM 3D stack configuration in section 7.1.

Stress sensors, such as the devices analyzed in chapters 5 and 6 are used to translate current shift to mechanical stress. Finite element modeling can be used to simulate the mechanical stress and translate it to current shift. Such a task was performed in figure 7.3 a) where current shifts obtained with finite element modeling were compared to current shifts from electrical measurements. Following eq. 3.10 and 3.11, the stress components are obtained through finite element modeling and transfered to current shift via the experimentally obtained piezocoefficients, as in chapter 5, specific to the device of interest.

The current shift on the top side of the thinned Si die opposite of the microbump is governed by the two dominant in-plane stress components, as seen in figures 7.3 and 7.4. The occuring current shift patterns in n-type and p-type Si from the underfill-microbump mechanism on the opposite side of the microbump are simulated and presented in figure 7.9. The projected contours of the underlying microbump are plotted on the current shift patterns with dashed circular lines.

N-type Si exhibits circular positive current shift patterns with the highest value

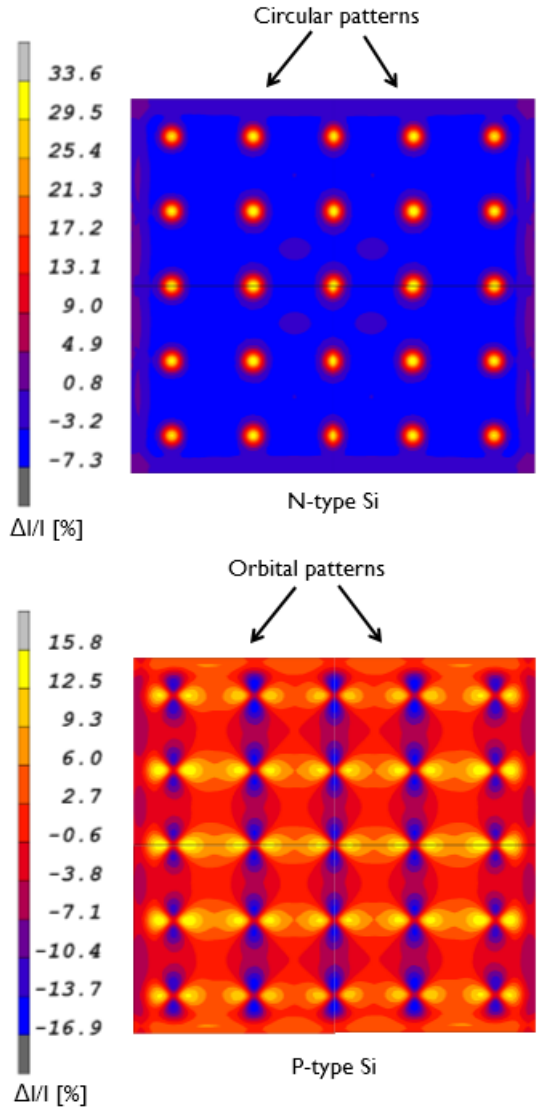


Figure 7.10: Simulated n-type and p-type current shift patterns on the top thinned Si side above a 5x5 underlying microbump array. When the same underfill from figures 7.4 to 7.6 is used, a current shift of +33% is expected in the peaks of the n-type Si circular regions and about 16%, in positive and negative current shift orbitals for p-type Si. Placing a transistor closer or further from the microbump and in which area can make all the difference in its performance.

directly above the microbump fading towards the edges and beyond. Further away from the microbump, the sole underfill impact region is entered and the current shifts change to slightly negative. P-type Si exhibits mixed positive and negative orbital current shifts with the lowest value above the microbump and the highest close to the edges of the microbump. With distance, the current shift in p-type Si fades to zero. These current shift patterns are a result of fundamental n-type and p-type Si response to in-plane stresses in [110] and [-110] direction, respectively and the basic shapes do not change with changing geometry, material properties of the 3D stack or sensitivities of the transistors. The absolute current shift values and reach of the current shift patterns are the ones that do change. The current shift patterns from figure 7.9 imply what effect a n-type or p-type transistor will feel if placed in these regions.

The n-type and p-type current shift patterns are a consequence of the interaction of the signs of the piezocoefficients and occuring stresses. When neglecting out-of-plane stress, as can be done in this case according to figure 7.4 and looking into eq. 5.7 and 5.8 and stress sensitivities in table 5.1, the following can be stated for n-type and p-type current shift shapes:

- N-type: The tensile stresses occuring in Si, σ_l and σ_t, going as positive into the current shift eq. 5.7 or 5.8, are multiplied by piezocoefficients π_l and π_t which are both negative. The addition of $\pi_l\sigma_l + \pi_t\sigma_t$ is constructive resulting in the highest value in the center fading equally in all directions from the center of the microbump forming a circular shape

- P-type: The tensile stresses occuring in Si, σ_l and σ_t, going as positive into the current shift eq. 5.7 or 5.8, are multiplied by piezocoefficients π_l and π_t which are of opposite signs. The addition of $\pi_l\sigma_l + \pi_t\sigma_t$ is destructive and contributions of the two in-plane components work against each other in the superposition lowering the final current shift value. Since the two in-plane piezocoefficients, π_l and π_t are of similar absolute value and the in-plane stresses σ_l and σ_t are equal directly above the microbump, the total current shift directly above the microbump is close to zero. The orbital current shift shapes start appearing further away from the microbump where one in-plane stress starts to be dominant and over its corresponding positive or negative piezocoefficient pushes the current shift into negative or positive values, respectively.

Figure 7.10 plots n-type and p-type current shift patterns on the top thinned Si side above a 5 x 5 underlying microbump array. When the same underfill from figures 7.4 to 7.6 is used, a current shift of +33% is expected in the peaks of the n-type Si circular regions and about 16%, in positive and negative current shift orbitals for p-type Si. Placing a transitor closer or further from the microbump and in which area can make all the difference in its performance.

Potential keep-out zones are proposed for n-type and p-type Si in figures 7.11 and 7.12, respectively. Figure 7.11 presents a study on n-type Si. Close to +30% of current shift is observed in Si on the opposite side of the microbump. The current shift drops to about -4% 60 μm away from the microbump and further on settles at a value of -1%. If a rule is set for a n-type transistor that its current shift should not be affected more than 2%, there are certain regions the transistor should not be placed in. The keep-out

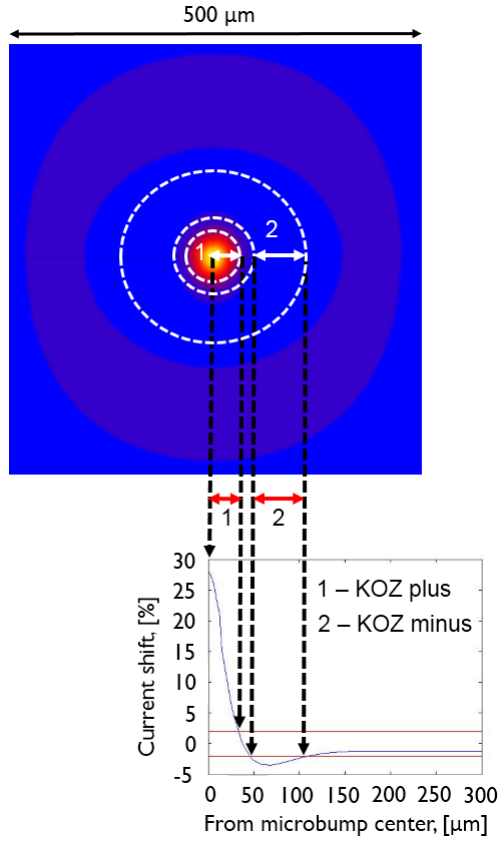


Figure 7.11: A case study including potential keep-out zones for n-type Si with a design rule of 2% allowed current shift

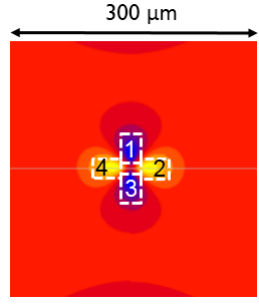


Figure 7.12: Potential keep-out zones p-type Si

Parameter	Underfill	Thick Si die thickness	Thin Si die thickness	Microbump pitch	Microbump height
Initial value	UF 6-A High stress	700 µm	25 µm	200 µm	13 µm

Table 7.2: 3D IC stack properties of the starting finite element model based on the logic on DRAM stack

regions equals to the areas with current shift higher than +2% and lower than -2%. This means the transistor should not be placed in regions from +2% to +30%, indicated as region one and from -2% to -4%, indicated as region 2. For n-type Si, on the Si surface these keep-out regions are belts. In conclusion, for n-type Si, the keep-out zones are circular, in this case consisting of two belts where the transistor should no be placed during IC layout production.

For p-type Si, figure 7.12, the keep-out zones are more complicated as the current shift shapes are more complex. Depending on the allowed current shift for a transistor, 4 keep-out zones can be defined around the current shift orbitals, circular, or in a safer case, rectangular. These keep-out zones will not be present directly above the microbump but from the microbump edges onwards.

In order to compare FEOL stress fields originating from a TSV and the underfill-microbump interaction, in figure 7.13 impact zones from a TSV and microbump are compared. Stress in Si from a 5μm x 50μm TSV and a 30 μm in diameter microbump is simulated and compared. Underfill properties were delibarately altered by varying its CTE to approximately match the absolute stress peaks of the TSV. Both the TSV and the underfill-microbump mechanism now exhibit a maximum of close to 300 MPa on the top side of the thinned die. However, looking at the spread of stresses around the microbump and TSV, the impact zone of the microbump is approximately 10 times larger than the impact zone of a TSV. This implies that although both the underfill-microbump mechanism and the TSV can create concerning stresses in the FEOL, the underfill-microbump mechanism can affect an area much larger than the one of a TSV. Consequently, a microbump keep-out zone can be considerably larger than a TSV keep-out zone.

7.3 Underfill-microbump stress mitigation guidelines

7.3.1 Parametric study

A parametric study based on an experimentally validated finite element model was employed in order to form guidelines for stress reduction caused by the underfill-microbump mechanism. The logic on DRAM finite element model was compared with electrical measurements, figure 7.3 a), and was therefore taken as the starting point for the parametric study. The following stack properties were investigated:

- material based

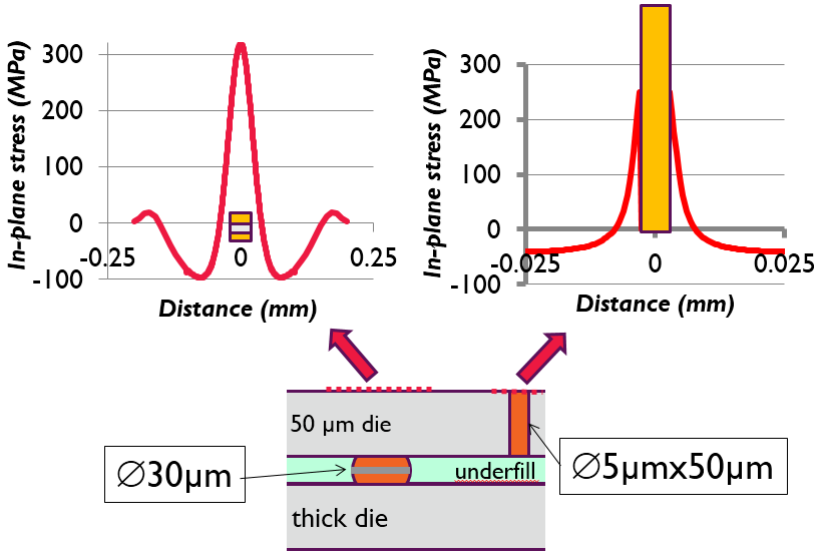


Figure 7.13: Simulated impact zones from a TSV and microbump. In case of equal absolute stresses generated on the top Si side opposite of the microbump, the impact zone of the microbump is approximately 10 times larger than the impact zone of a TSV.

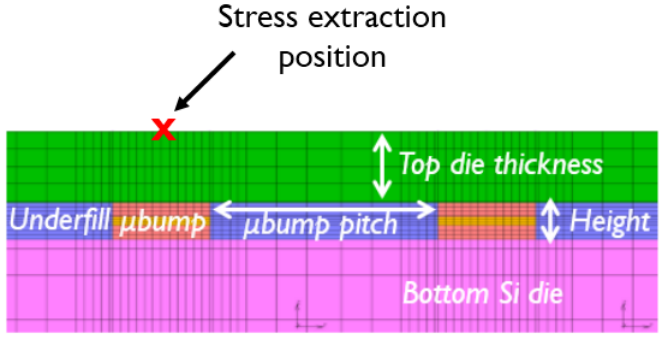


Figure 7.14: Illustration pinpointing the varied parameters and the stress extraction position in the finite element model of the logic on DRAM stack

  - choice of underfill

- stack geometry based

  - Si die thickness
  - microbump pitch
  - microbump height

The parametric study also includes observation of a microbump array edge effect and stress reduction by introduction of novel 3D stack features, Si etch rings.

The initial properties of the starting finite element model based on the logic on DRAM stack are summarized in table 7.2. An illustration pinpointing the varied parameters and the stress extraction position is shown in figure 7.14. The stress on top of the thin Si die opposite of the microbump was extracted as major principal stress, which is equal to the two in-plane stress components on the Si surface. Each new simulation included variation of only one parameter from the initial values to investigate their individual contributions. All underfills are coded due to material supplier requests. Si die thickness refers to the thickness of the top die submitted to thinning. In this case a thin die is always stacked on a thick 700 μm die. Although the original logic on DRAM stack included a thick die of 550 μm, the finite element simulation used to represent the logic on DRAM stack was performed with a thickness of 700 μm as the precise thickness was known only after stack processing and electrical measurements. Finite element simulations conducted later confirmed no visible difference in impact to the thin die for both thick die thicknesses. Both a 550 μm and 700 μm thick die act as a rigid body resulting in the shrinking underfill causing warpages only to the thin die. The microbump pitch is defined as the distance from one microbump center to the other neighbouring microbump center. The initial microbump height of 13 μm includes two Cu pads of 5 μm each and 3 μm of Cu-Sn intermetallic in between.

7.3.2 Underfill selection

Following results from sections 7.1 and 7.2, it is conclusive that the presence of underfill ignites an impact on the FEOL that presents itself in full form in the vicinity of microbumps. The logic on DRAM stack with initial underfill 6-A exhibited over 800 MPa of stress. In order to select an underfill for further stack prototypes that induces less stress, a series of underfills were simulated, presented in figure 7.15. The underfill names are coded due to a non-disclosure agreement with the material suppliers.

The underfill properties for the finite element simulations were obtained through DMA and TMA curves given by the material supplier or through datasheet values where full curves were not available. Each underfill was represented with its CTE and Young's modulus above and below its glass transition temperature and cure shrinkage. Most underfills exhibited a stable value of CTE and Young's modulus above and below the glass transition temperature. Approximately 10 ºC above and below the glass transition temperature defined the underfills' transition region. Through this region counting down from bonding temperature, which is above T_g, to room temperature, the

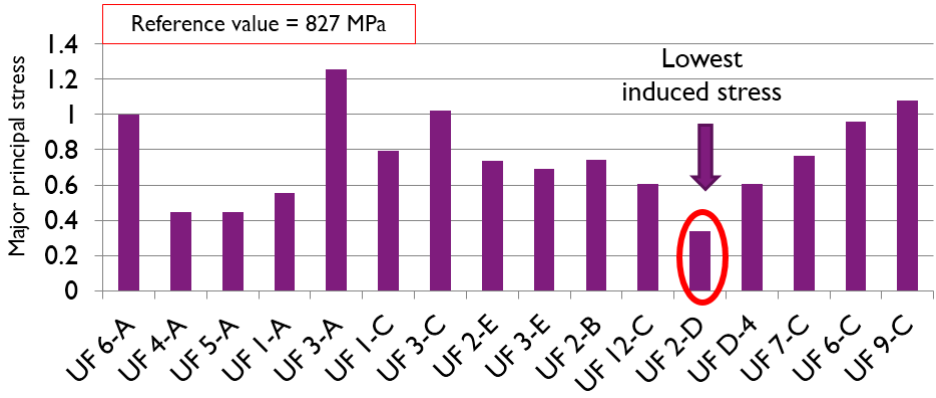


Figure 7.15: A series of simulated underfills from various material suppliers compared by the stress they generated in Si

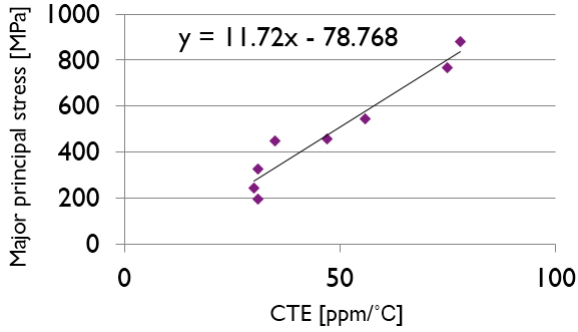


Figure 7.16: Stress in Si versus CTE values below the glass transition temperature of the first 8 underfills from 7.15

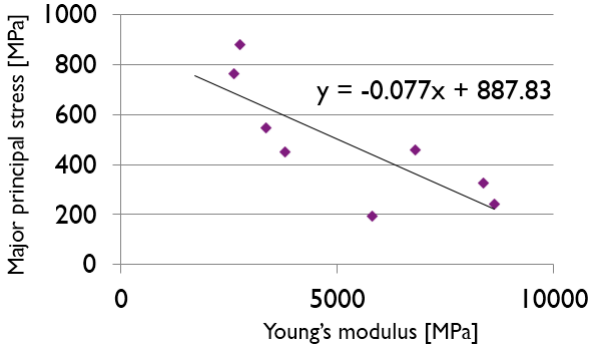


Figure 7.17: Stress in Si versus Young's modulus values below the glass transition temperature of the first 8 underfills from 7.15

underfill solidifies and the initially low Young's modulus significantly increases by an order of magnitude and the CTE decreases to 2-3 times from its initial value.

Figure 7.15 presents a variety of caused stresses in Si. It is indicative that from a thermo-mechanical point of view, several underfills with less impact on Si than UF 6-A are available. In particular, from this group of simulated underfills, UF 2-D causes the least stress in Si, in total 3 times less than UF 6-A. Therefore, the choice of underfill should be based on not just its chemical performance or application method but its thermo-mechanical impact is of high importance as it can significantly reduce stress levels in Si.

Figure 7.16 plots the CTE values below the glass transition temperature of the first 8 underfills from figure 7.15 versus their stress impact. The stress values are slightly lower than in figure 7.15 as these simulations were conducted without underfill cure shrinkage considered. The underfill CTE has a significant impact on Si stress levels. According to figure 7.16 an underfill with 10 ppm/ºC higher CTE can cause more than 100 MPa higher stress. A similar plot was made with the Young's modulus of the first 8 underfills, presented in figure 7.17. The curve would suggest that a higher Young's modulus would imply lower stress. From a fundmental viewpoint, since the underfill CTE is much higher than the one of Si and since the basis of the underfill-microbump mechanism is the underfill pulling on the Si while shrinking, it is more logical to expect that a higher Young's modulus of underfill causes higher stress in Si. Looking into the processing of underfill, figure 7.17 can be interpreted in a different manner. The basic components of standard underfill for microelectronics is the epoxy material with filler particles. In general, adding filler particles, such as SiO_2, to epoxy lowers the CTE of the underfill and increases the Young's modulus. Therefore, the CTE and Young's modulus of underfill is reversly proportional. The underfills with high CTE in figure 7.16 are in fact the ones with a lower Young's modulus in figure 7.17. To assess the reverse proportionality of Young's modulus and CTE, in figure 7.18 an inverse of Young's modulus versus the CTE of the first 8 underfills is plotted, according to eq. 7.1. E stands for the Young's modulus. The values deviate slightly from a purely linear relationship, but the trend of lower Young's modulus with higher CTE is observed. Therefore, the CTE rather than the Young's modulus bears a higher importance in the underfills thermo-mechanical performance.

$$E \cdot CTE = const. \tag{7.1}$$

The glass transition temperature, T_g, acts as the dividing temperature of two regions, the soft underfill with higher CTE above T_g and harder underfill with lower CTE below T_g. The region from bonding temperature to glass transition temperature usually around 100 ºC presents one phase of stress build up and the transition from glass transition temperature to room temperature another stress build up phase. Expanding and decreasing these regions by shifting the T_g can cause a difference in final accumulated stress at room temperature. Figure 7.19 presents results of T_g variation of underfill UF 6-A and the resulting stress shift in Si. Although the impact of underfill's CTE is by far greater than the one of T_g, stress shifts are observed with approximately 5% lower stress with 10 ºC higher T_g. It can be discussed to which extent the underfill can generate stress above its glass transition temperature considering its very soft nature. The

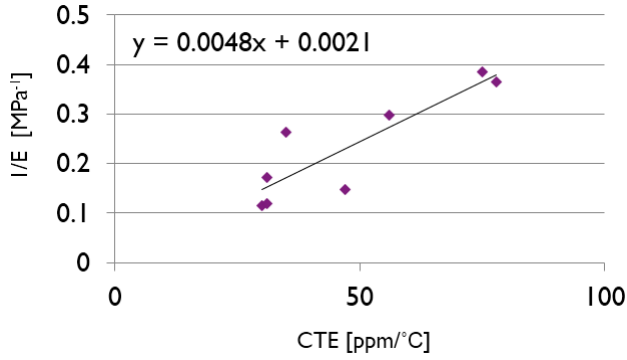


Figure 7.18: An inverse of Young's modulus versus the CTE of the first 8 underfills is plotted, in support of eq. 7.1

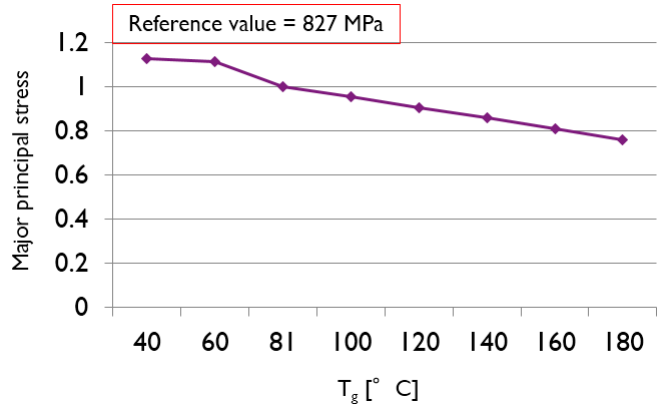


Figure 7.19: Stress in Si versus T_g variation of underfill UF 6-A

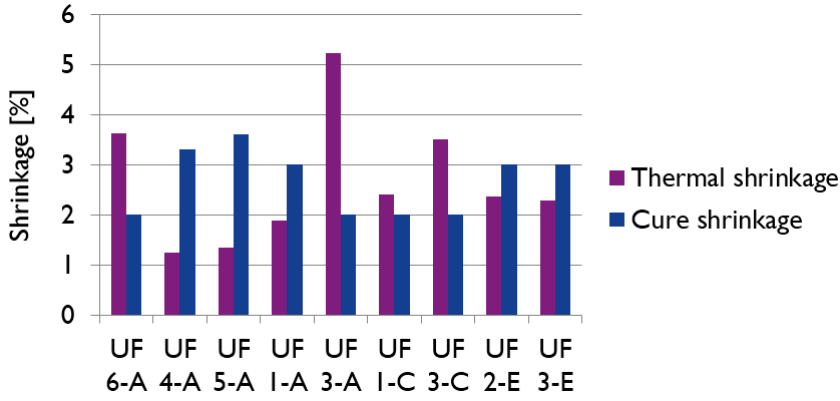


Figure 7.20: The thermal shrinkages and cure shrinkages of the first 8 underfills

generation of stress at higher temperatures is further investigated in section 7.4.2.

A faster approach for comparing stress impact from a series of underfills is proposed by calculating underfill thermal shrinkages. Underfill thermal shrinkage, TS, is given as

$$TS = (250\text{ºC} - T_g) \cdot CTE_{above} + (T_g - 25\text{ºC}) \cdot CTE_{below} \tag{7.2}$$

where CTE_{above} stands for CTE above the glass transition temperature and CTE_{below} stands for CTE below the glass transition temperature. Thermal shrinkage from eq. 7.2, TS, is a sum of material shrinkage in percentage above and below the glass transition temperature considering constant CTE values in the two regions. Apart from thermal shrinkage, it is considered that the underfill cure shrinkage plays a role in stress generation as well. The thermal shrinkages and cure shrinkages of the first 8 underfills are shown in figure 7.20. Both thermal and cure shrinkage vary between 1% and 5%, however thermal shrinkage has a significantly higher impact as it acts through two regions and changing underfill properties while the cure shrinkage occurs from underfill viscous liquid to solid phase.

Figure 7.21 plots thermal shrinkages of the first 8 underfills versus finite element obtained stress in Si. Their dependency is fairly linear. If one wishes to consider a new underfill, in first approximation, its CTE values above and below T_g can be taken from datasheets to calculate its thermal shrinkage and correlate this to stress using the generated linear equation. Although several initial finite element simulations still need to be made in order to establish a linear curve, afterwards other underfills can be in first approximation compared to simulated underfills in a fast manner by calculating their thermal shrinkages.

The impact of cure shrinkage in simulations was added by artificially raising the bonding temperature depending on the cure shrinkage value according to eq. 7.3, where ε_C stands for the cure shrinkage, CTE_{above} the CTE above the glass transition temperature and $\triangle T$ the temperature which is added to the bonding temperature of 250 ºC.

$$\triangle T = \frac{\varepsilon_C}{3 \cdot CTE_{above}} \tag{7.3}$$

Figure 7.22 presents simulation results with varying cure shrinkage. According to simulations and a simplistic inclusion of cure shrinkage using eq. 7.3, 1% decrease of cure shrinkage results in approximately 5% stress decrease in Si. It is believed that cure shrinkage has a minor but non-negligible impact on stress generation and should be considered in the underfill-microbump mechanism, however more fundamental incorporation of the development and impact of cure shrinkage was out of scope of this thesis.

7.3.3 Impact of stack geometry

Apart from underfill material properties, stack geometry was addressed in pursuit of stress reductions in Si. Since Si becomes flexible with thinning and prone to bending from the shrinking underfill, increasing Si thickness offers itself as the first solution.

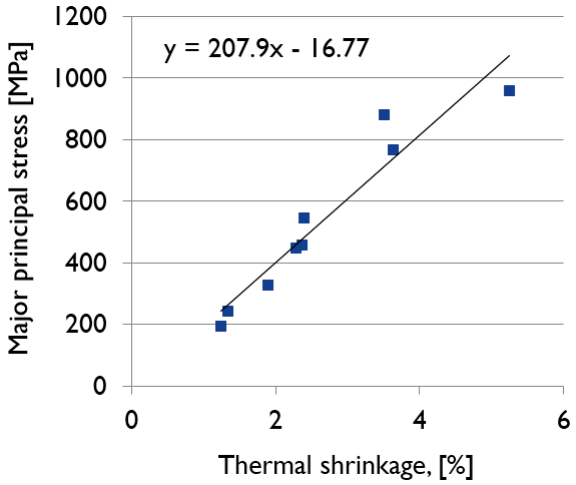


Figure 7.21: Thermal shrinkages of the first 8 underfills versus finite element obtained stress in Si

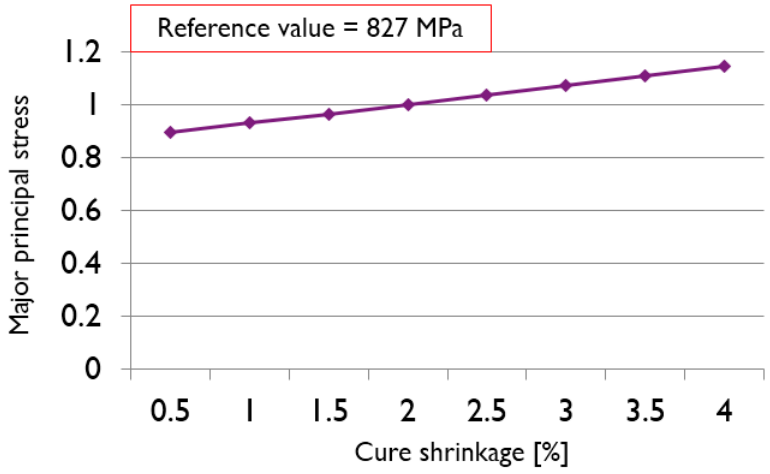


Figure 7.22: Stress in Si versus cure shrinkage variation of underfill UF 6-A

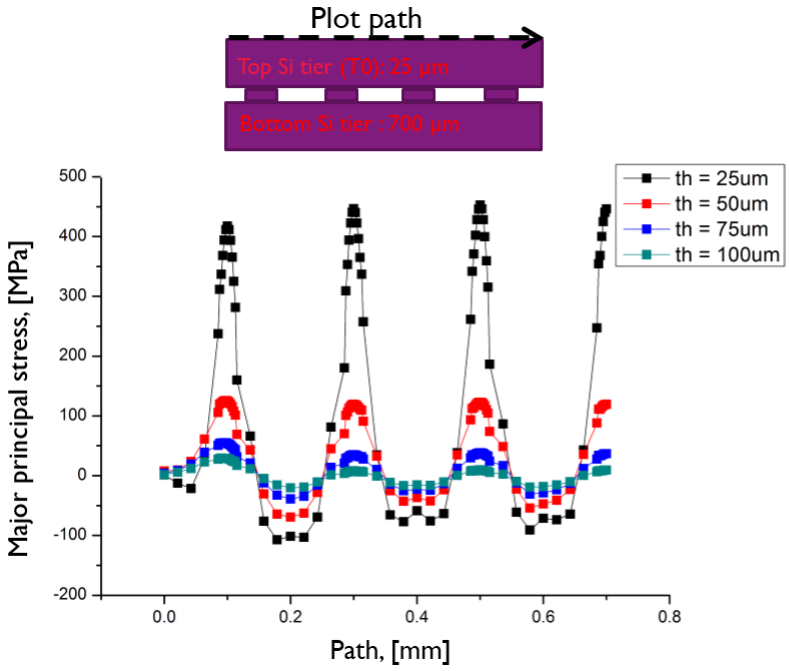


Figure 7.23: Stress on the top Si surface with underfill UF 4-A over a path involving 4 underlying microbumps. Si thickness is varied. The oval stress pattern above each microbump corresponds to the current shift oval pattern observed in figure 7.3 a).

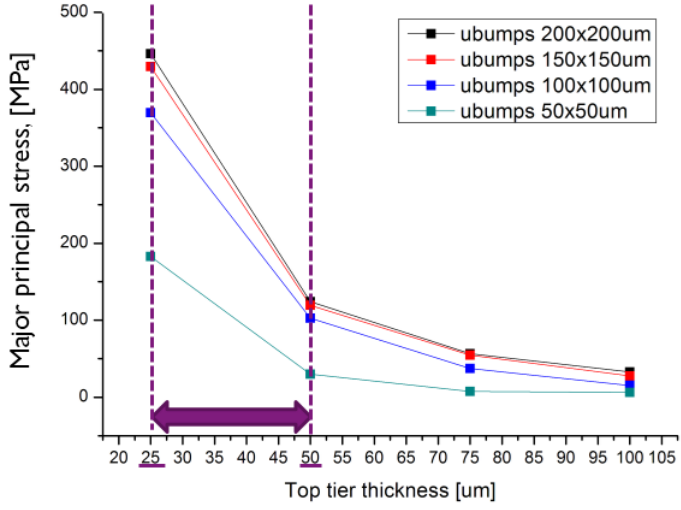


Figure 7.24: Effect of top Si thickness and microbump pitch on stress on the top Si surface opposite of the microbump

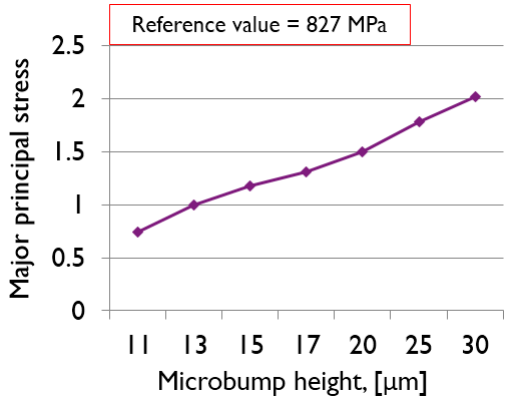


Figure 7.25: Effect of microbump height on stress on the top Si surface opposite of the microbump

Figure 7.23 plots stress on the top Si surface with underfill UF 4-A over a path involving 4 underlying microbumps versus the top Si thickness. The oval stress pattern above each microbump corresponds to the current shift oval pattern observed in figure 7.3 a). The wave stress pattern above several microbumps diminishes when top Si thickness is increased from 25 μm. At 100 μm the stress is below 50 MPa, 8 times lower than an initial value of 400 MPa.

Figure 7.24 plots the effect of top Si thickness and microbump pitch. The initial microbump pitch of 200 μm when microbump stress impact zones are not overlapping is lowered to 50 μm. Already at a pitch of 100 μm stress is visibly decreasing, while for 50 μm the stress is 2.5 times lower. When microbumps are placed closer together within each others impact zones, they start supporting the thin top Si preventing it from bending. A standalone microbump will induce more stress to Si than when microbumps are grouped within a sufficiently short pitched array. A sole increase of Si thickness from 25 μm to 50 μm will cause 3 times lower stress while additionaly lowering microbump pitch to in this case 50 μm will lower stress 8 times, the same as in figure 7.23 with just increasing the thickness to 100 μm. Therefore, a clever design of microbumps along with increasing Si thickness can significantly reduce underfill-microbump generated stress.

Figure 7.25 presents results of finite element simulations when the original microbump height of 13 μm is varied. Increasing the microbump height implies more underfill material, leading to considerably higher stress predictions. A decrease from an intial height of 13 μm to 11 μm causes 25 % stress decrease. When closing the gap below 10 μm, attention needs to be paid to controling filler particle sizes.

While closing the microbump pitch reduces Si stress above the microbump array, another effect seems to appear on the periphery of the array. Figure 7.26 presents the microbump array edge effect. When microbumps are placed in sufficient vicinity to prevent bending of the thinned Si die, creating a microbump array, the microbumps within the array indeed exhibit a reduction of stress. However, Si above the mi-

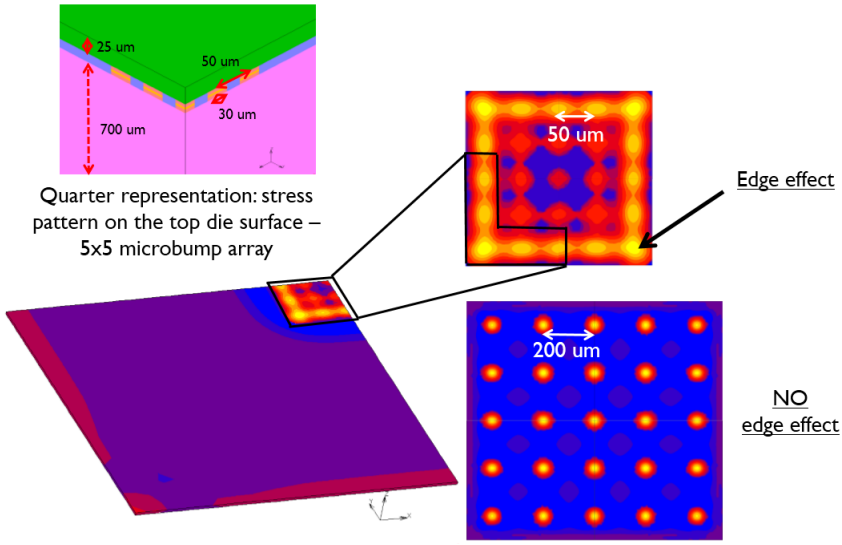


Figure 7.26: The microbump array edge effect. Microbumps placed in an array at sufficient proximity cause less stress in the Si above within the array however Si above the microbumps forming the edge of the microbump array exhibits stress peaks.

crobumps forming the edge of the microbump array exhibits stress peaks. Although locally within the array, Si is supported by the presence of the microbumps, the microbump array as a whole acts as a platform over which underfill during shrinking pulls the Si over its edges.

Figure 7.27 provides more explanation on the evolution of the microbump edge effect. A standalone microbump is gradually increased to a 5x5 microbump array while monitoring the stress on the microbump in the middle of the array and on its edges. A 5x5 microbump array with large 200 μm pitch exhibits almost the same stress as an individual microbump as with the larger pitch their impact zones barely touch each other and there is no formation of support for Si. There is no stress reduction when another microbump is added at 50 μm distance. The first stress reduction is visible when a 2x2 array with 50 μm pitch is formed. All 4 microbumps are now edge microbumps and the edge effect is dominant compared to the contribution of Si support. The two effects, firstly the supporting action of microbumps below Si and secondly the edge effect can be firstly distinguished on a 3x3 size array where separately an inner microbump and edge microbumps are present. Si above the middle microbumps exhibits stress reduction because the microbump array locally prevents the top Si die from bending while the edge effect contributes to the stress increase above the outer microbumps. With increase of the array, the microbumps on the edges still induce the same stress. The inner microbumps exhibit additional stress reduction with increasing the microbump array size. Therefore, when grouping microbumps within an array close enough that their stress fields start overlapping, lowering the microbump pitch and increasing the microbump array size will contribute to the stress decrease in Si within the microbump

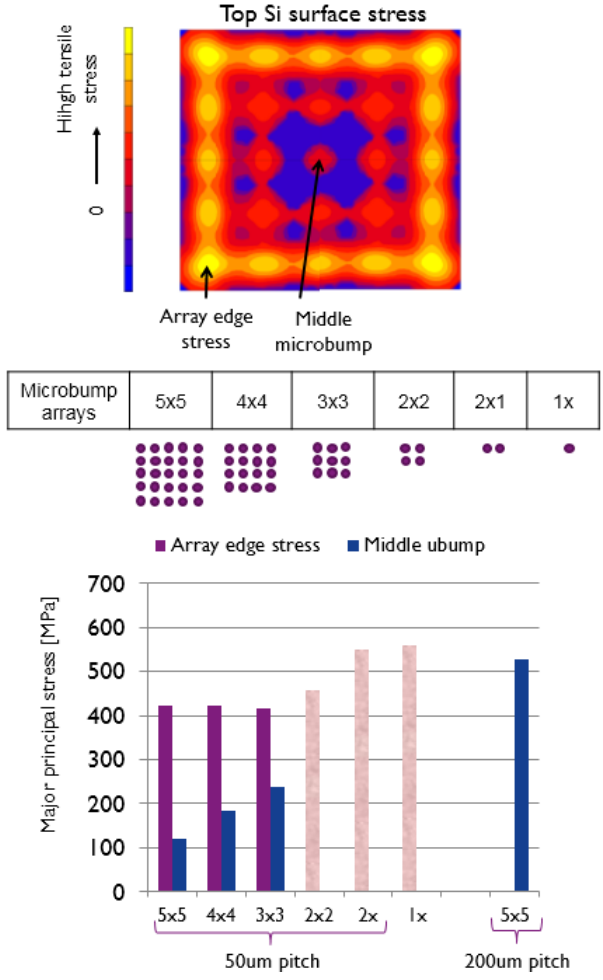


Figure 7.27: Evolution of the microbump edge effect. With increasing the microbump array size, stress within the array reduces, the edge stress remains after forming edges on a 3x3 array remains the same.

array while stress peaks can be expected in Si on the edges of the microbump array due to the microbump array edge effect.

7.3.4 Si etch ring

A conceptual study involving introduction of additional material on the microbump side and etched filled rings in Si in order to reduce underfill-microbump generated stress was filed as a patent, publication number US 2013/0154112. The concept was supported solely with finite element modeling. No electrical measurements or any other tests have been conducted to this point to prove the stress reduction procedure. The patent served as protection of a methodology.

Figure 7.28 illustrates the proposed stress reduction procedure in several steps with figure 7.29 presenting the corresponding simulated stress values in Si on the opposite side of the microbump as in figure 7.14. Table 7.3 summarizes the material properties used in the finite element model. All Poisson's ratios were set to 0.3. Cu plasticity was included by a piecewise linear function, an idealized work hardening stress/strain curve, presented in figure 4.4. In figure 7.28 a), the starting point is a 2-layer stack, with thin Si stacked on top of thick over Cu-Sn microbumps. This case corresponds to the underfill-microbump mechanism and observed current shifts in figure 7.3. Additional layers of material such as SiN, figure 7.28 b), or Benzocyclobutene (BCB), figure 7.28 c) can be added in the first step to decrease stress. The basic idea is to add thin layers of material between the underfill and Si, with a Young's modulus lower than Si. Lower amounts of stress can be observed in Si for the following reasons:

- The shrinking underfill transfers its displacements during shrinking over both the newly added material and Si. Ultimately, the displacements transferred to Si are less. This implies less strain and therefore less stress in Si.
- If the same microbump height is maintained, adding material between two Si dies means decreasing the amount of underfill material. As the thermal shrinkage of the underfill remains the same, a smaller thickness of the underfill means also less underfill displacement that will be transfered to surrounding materials in the first place.
- Adding materials between the underfill and Si changes the equivalent Young's modulus of the material system that the underfill is pulling on. A material with lower Young's modulus than Si causes the Si-added material system to exhibit lower equivalent Young's modulus. In result, in reaction to the displacement coming from the underfill during shrinking less stress is transfered to Si than without the added material.

The first column in figure 7.29 reflects the stress at the starting point, figure 7.28 a). When 1.5 μm of SiN is added, second column of figure 7.29 and illustration in figure 7.28 b), stress slightly drops. If the same thickness of BCB is now added with significantly lower Young's modulus than Si, third column in 7.29, stress drops even more. A thicker BCB adds an additional stress reduction, 4th column figure 7.29. A combination can be made with 3 μm of BCB placed on 1.5 μm of SiN, figure 7.28 d) which

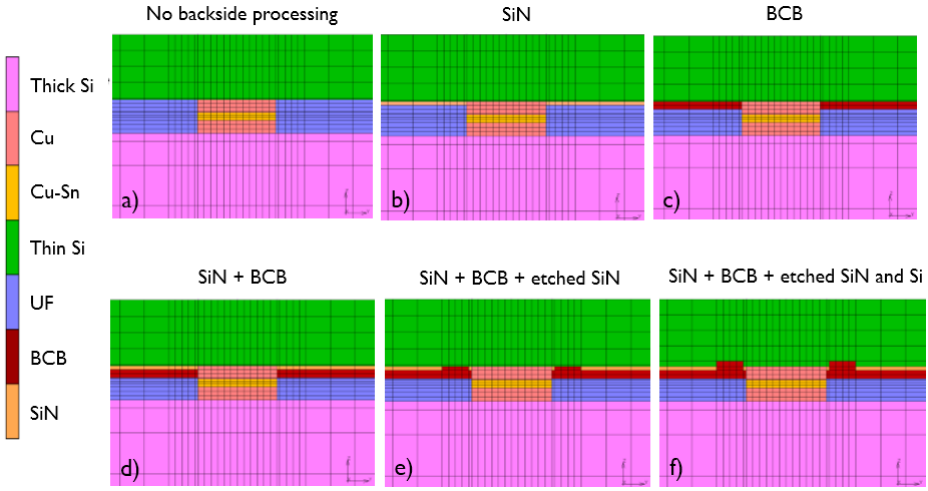


Figure 7.28: Illustration of the proposed Si etch ring stress reduction procedure in several steps

results in 15 % less stress in Si, figure 7.29 5th column. More stress is generated closer to the microbump, so in attempt to lower the equivalent Young's modulus closer to Si, SiN can be etched and filled with BCB, figure 7.28 e). Ultimately, Si itself can also be etched and filled with BCB. These formations are circular creating rings of material in etched SiN and Si around the microbump. Etching SiN and 0.5 μm of Si cause an additional 2% drop of stress in Si, figure 7.29 7th column. In this hypothetical case, FEOL devices could not be processed in Si etch defined regions.

The width and depth of these rings and their proximity to the microbump can be varied to optimize their influence. The closer they are to the microbump the closer they will be to higher stresses needed to be lowered. The wider they are, the larger area of the underfill-microbump impact zone they will cover. With ring depth, equivalent Young's modulus futher lowers, however Si becomes also locally more flexible. Ring depths higher than 2 μm on 25 μm in total thick Si showed again trends of stress increase to initial values, last 3 columns of figure 7.29.

7.3.5 Ranking of stress reduction parameters

Incoropration of new geometrical and material aspects of 3D IC stacks simulated in sections from 7.3.2 to 7.3.4 can significantly decrease Si stress making strong steps towards resolving the impact of the underfill-microbump stress mechanism. A case study was performed in order to rank the simulated parameters by their stress reduction impact. The parameters of the initial logic on DRAM stack and the modified stack are presented in table 7.4. The following changes have been made to the new stack:

- High stress underfill UF 6-A was replaced by lower stress underfill UF 1-A

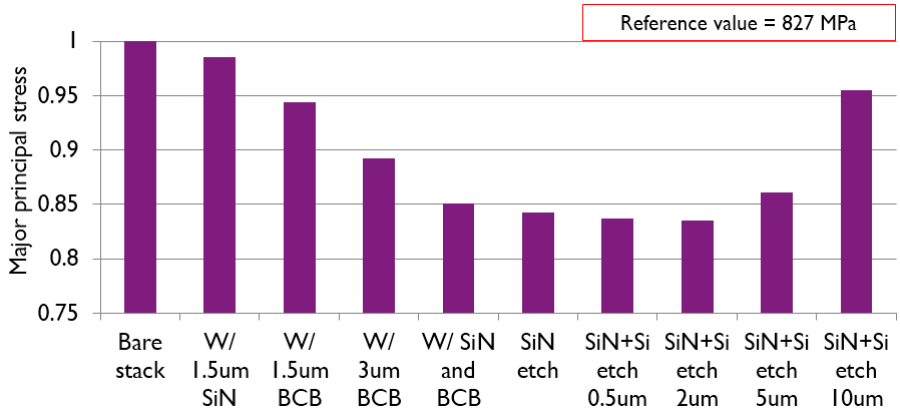


Figure 7.29: Simulated stress values in Si on the opposite side of the microbump corresponding to steps in figure 7.28

Material properties	Si	SiN	BCB	Cu	Cu-Sn	Underfill
Young's modulus [GPa]	169	90	2.9	117	108	$T<T_g$ 2-8 $T>T_g$ 0.1
CTE [ppm/MPa]	2.6	3.3	42	16	19	$T<T_g$ 30-80 $T>T_g$ 100+

Table 7.3: Material properties used in the Si etch ring finite element model

Case study	Underfill	Etch rings	Thin die thickness	Microbump pitch	Microbump height	Microbump array
DRAM on logic stack	UF 6-A	None	25 µm	200 µm	13 µm	Standalone microbumps
Modified stack	UF 1-A	1.5 µm SiN + 3 µm BCB + 0.5 µm ring in Si	100 µm	50 µm	11 µm	5x5 array

Table 7.4: Parameters of the initial logic on DRAM stack and the modified stack

3D IC stack parameter	Underfill CTE	Underfill T_g	Underfill Cure shrinkage	Si die thickness	Microbump pitch	Microbump height	Microbump array size
Trend for stress decrease	⬇	⬆	⬇	⬆	⬇	⬇	⬆

Table 7.5: A summary of guidelines for underfill-microbump stress mitigation involving 3D IC stack assembly trends for stress reduction

- Etch rings were added: 1.5 μm of SiN, 3 μm of BCB and BCB filled rings through SiN and 0.5 μm of Si
- Top thinned Si thickness was increased from 25 μm to 100 μm
- Pitch between microbumps was lowered from 200 μm to 50 μm
- Microbump height was decreased from 13 μm to 11 μm
- Array size, not just microbump pitch was taken into consideration by creating a 5x5 microbump array

Figure 7.30 presents the stress decrease contributions of each of these modifications and ranks them according to benefiting impact. The highest stress reduction came from the increase of Si thickness with 5 times lower stress. Si thickness was followed by a close to 60% stress drop, equally from usage of low stress underfill or lowering the microbump pitch. Microbump height and microbump array size contributed equally with approximately 25% stress decrease and adding etch rings ends the ranking with a 17% stress drop.

All properties were included in one model to obtain the combinatory effect of all of the mentioned parameters. The new stack with all incorporated modifications now exhibited 10% of the original stress in Si. A summary of guidelines for underfill-microbump stress mitigation involving 3D IC stack assembly trends for stress reduction is presented in table 7.5. Etch ring is excluded from the summary, as a new feature that can be added to existing designs.

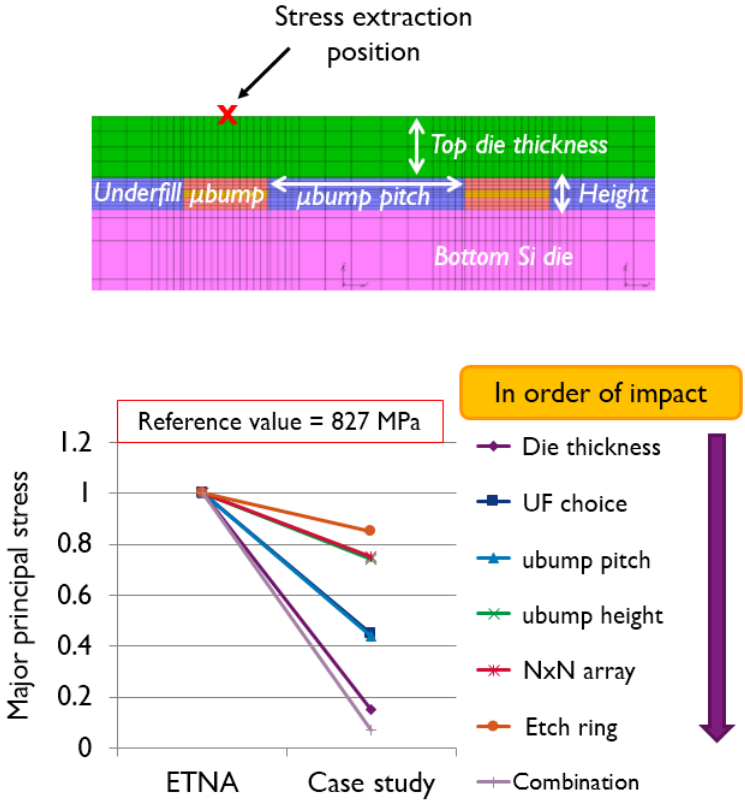


Figure 7.30: Stress decrease contributions of stack modifications simulated in sections 7.3.1 to 7.3.4 and their ranking according to benefiting impact

7.4 Investigation of stress mitigation on test chips

7.4.1 Test chips for 3D IC stacking

The first impact of the underfill-microbump stress mechanism was observed on the logic on DRAM 3D stack, where the logic die consisted of the 130 nm ETNA test chip. Several test chips followed the ETNA test chip., 3D 65, FUJI and PTCQ. 3D 65 test chip, was processed only for assessment of 65 nm FEOL logic devices and was not stacked or packaged. FUJI was used for 32 nm FEOL stress calibration and submitted to 3D stacking. PTCQ is a dedicated stress test chip processed in 65 nm technology which incorporated learnings from forementioned test chips. It comprised FEOL and BEOL integrated stress sensors developed primarily for assessment of 3D stacking and 3D packaging.

7.4.2 PTCQ stress test chip design

The PTCQ stress chip, figure 7.31, with width and length 8 mm x 8 mm is organized into 16 basic modules arranged in a 4x4 fashion. Each basic module, 1.92 mm x 1.92 mm, consists of 64 cells, arranged in a 8x8 array. Each cell takes up an area of 240 μm x 240 μm. There are three types of cells, which distribution in each basic module is visible in figure 7.31. Cell 1 consists of stress sensors for monitoring global stress, cell 2 consists mainly of heaters and cell 3 consists of stress sensors for monitoring local stress. All I/O pads and control logic are present in the cross area in between the basic modules, as indicated in figure 7.32. The edges of the area also comprise wirebonding pads so the test chip is available for both flip chip and wirebond type 2D or 3D packaging.

Cell 1

Cell 1 consists of the following stress sensors used for monitoring global stress across the die:

- in-plane FEOL stress sensors
  - n-type and p-type MOSFETs, 4 μm x 4.4 μmm with current orientation in [100] and [010]
  - n-type and p-type Pseudo-Hall transistors, 3μm x 3 μm, with n-type transistor current in [110] and p-type current in [100]
- out-of-plane BEOL stress sensors
  - CIEF CBCM based capacitive sensors formed on metal levels 1 and 2, with capacitive plates of 30 μm x 30 μm on each of the two metal levels

The primary role of cell 1 is obtaining global stress values across the die in order to capture effects not visible on local area level. The cell also incorporates diodes for temperature measurements. Cu pads, 50 μm in diameter, 5 μm in height were placed

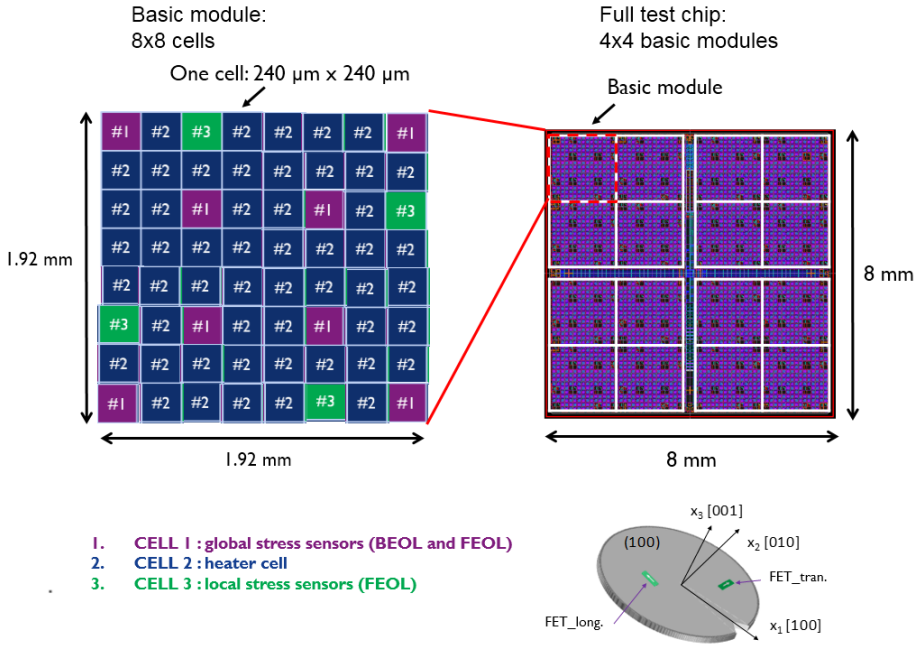


Figure 7.31: The PTCQ stress chip with width and length 8 mm x 8 mm is organized into 16 basic modules arranged in a 4x4 fashion. Each basic module, 1.92 mm x 1.92 mm, consists of 64 cells, arranged in a 8x8 array. Each cell takes up an area of 240 μm x 240 μm. There are three types of cells.

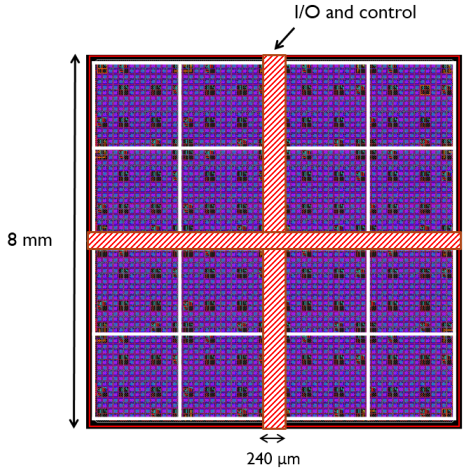


Figure 7.32: PTCQ I/O pads and control logic are present in the cross area in between the basic modules

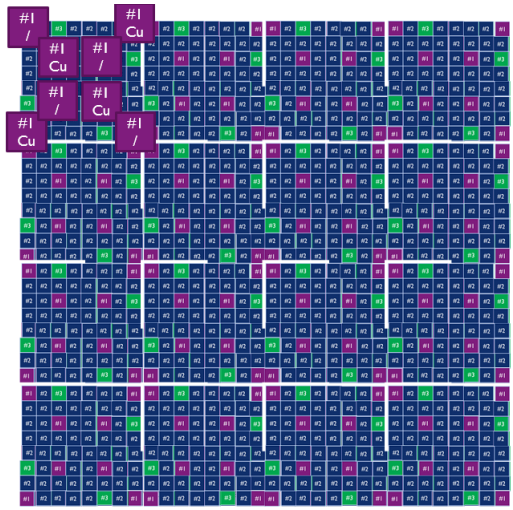


Figure 7.33: The Cu pad distribution above cells no. 1 within one basic module. 4 cells have a Cu pad on top, 4 are without a Cu pad. The Cu pad pattern was repeated on each of the 16 basic modules. Distribution of Cu pads only over the first basic model shown in the figure for better visibility.

on 4 out of 8 no. 1 cells within each basic module in order to monitor the impact of the Cu pad as well. The Cu pad distribution within one basic module is shown in figure 7.33. The Cu pad pattern was repeated on each of the 16 basic modules. Figure 7.33 shows the distribution of Cu pads over cell 1 positions only on the first basic module for better visibility.

Cell 2

Cell 2 consists of heating elements used to raise temperature to the die or parts of the die and also accompanies stress sensors used for monitoring global stress across the die:

- Heaters
  - 2 Cu plates processed on metal 1, 100 μm x 200 μm in size each
- in-plane FEOL stress sensors
  - n-type and p-type MOSFETs, 4 μm x 4.4 μmm with current orientation in [100] and [010]
  - n-type and p-type Pseudo-Hall transistors, 3μm x 3 μm, with n-type transistor current in [110] and p-type current in [100]

The primary role of cell 2 is global and local heat dissipation. The cell also incorporates diodes for temperature measurements.

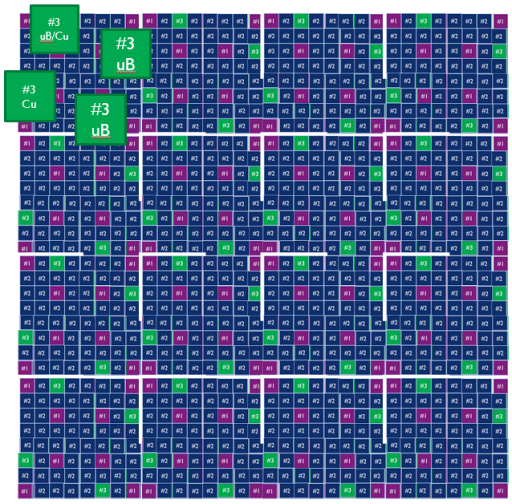


Figure 7.34: The Cu pad and microbump distribution over cells no. 3 within one basic module. The pattern is repeated on each of the 16 basic modules. Figure 7.34 shows the distribution of Cu pads and microbumps over cells no. 3 only on the first basic module for better visibility.

Cell 3

Cell 3 consists of the following stress sensors used for monitoring local stress in the die:

- in-plane FEOL stress sensors
  - n-type MOSFETs, 4 μm x 4.4 μm with current orientation in [100] and [010], arranged in a 7x8 transistor array

The primary role of cell 3 is assessing the underfill-microbump stress impact. The cell also incorporates diodes for temperature measurements. Cu pads, 50 μm in diameter, 5 μm in height were placed on 2 out of 4 no. 3 cells within each basic module. The Cu pad and microbump distribution over cells no. 3 within one basic module is shown in figure 7.34. The pattern was repeated on each of the 16 basic modules. Figure 7.34 shows the distribution of Cu pads and microbumps over cells no. 3 only on the first basic module for better visibility. Three combinations are now available, monitoring the underfill-microbump impact, monitoring the impact of the Cu pad and monitoring the combination of impacts from the underfill-microbump mechanism and Cu pad.

A detailed view of one no. 3 cell is presented in figure 7.35. One cell encompassed an array of 6x6 microbumps placed apart at a pitch of 40 μm. In case a cell with Cu pad presence, 50 μm in diameter Cu pads were placed above central microbumps in quadrants 2 and 4 of cell 3. The 7x8 nFET array was placed above the same central microbumps in quadrant 2 and quadrant 4 of the array. Additional individual nFETs were placed on the microbumps surrounding the central microbump and nFET array.

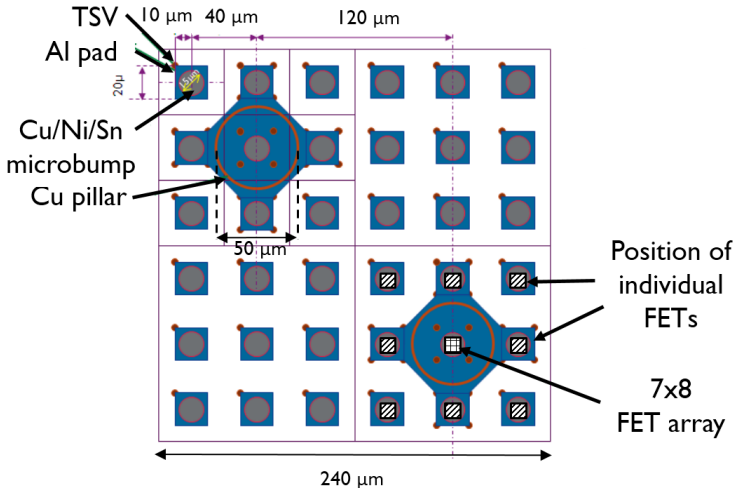


Figure 7.35: A detailed view of cell no. 3

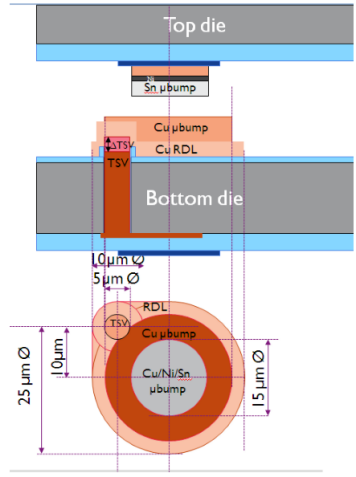


Figure 7.36: The Cu-Sn microbump on the PTCQ stack. The Cu pad on the thin die side is 25 μm in diameter, while on the thick die side the Cu pad and Sn are 15 μm in diameter.

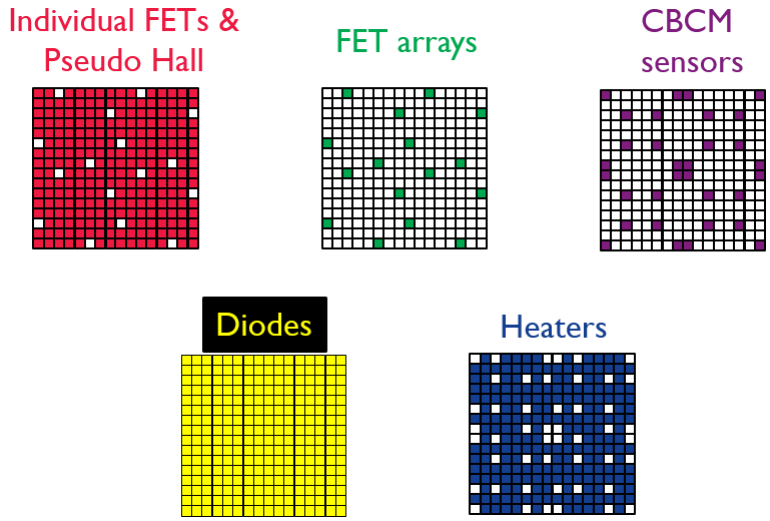


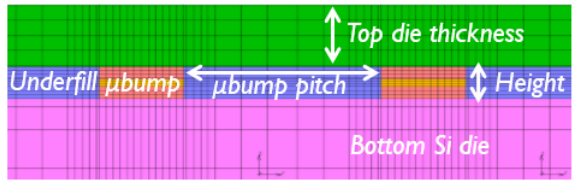
Figure 7.37: The coverage of all sensors over the die, including diodes and heaters

Figure 7.35 illustrates the nFETs only in quadrant 4 for better visibility although the same nFET pattern is present also in quadrant 2 of the array. The difference in the nFETs in quadrant 2 and quadrant 4 is that nFETs in quadrant 2 are oriented in [100] and nFETs in quadrant 4 in [010]. The Cu-Sn microbumps consisted of Cu pads with different dimensions, as indicated in figure 7.36. One side held a Cu and Sn pad of 15 μm in diameter while the Cu pad on the other side was 25 μm in diameter. The TSV was present on the Si side with larger Cu pads, connected on the side of the microbump, as shown in figure 7.36.

The coverage of all sensors over a quarter of the die, including diodes and heaters is shown in figure 7.37. Each square in figure 7.36 represents one cell, covering an area of 16x16 cells.

7.4.3 3D stack design comparison

The ETNA, FUJI and PTCQ test chips, sorted in chronological order, were used to produce 3D IC stacks. All produced stacks were 2-level stacks, i.e. a stack of two dies. The ETNA test chip was used as the logic die in a logic on DRAM configuration. FUJI dies were stacked on each other acting as a logic-on-logic test case. The PTCQ dies were also stacked on each other immitating a logic-on-logic test case, with stress sensors covered dies with a primary purpose of assessing stress levels after 3D assembly. Table 7.6 summarizes characteristic values of all 3 stacks generations.



3D IC stack	Die Technology node	Underfill	Bottom Si die thickness	Top Si die thickness	Microbump diameter	Microbump pitch	Microbump height	Microbump array
ETNA	130 nm	UF 6-A	550 µm	25 µm	30 µm	> 200 µm	13 µm	Standalone microbump
FUJI	32 nm	UF 6-A	700 µm	50 µm	30/15/25 µm	> 200 µm	13 µm	Standalone microbump
PTCQ	65 nm	UF 9-C	200 µm	50 µm	25/15/15 µm	40 µm	13 µm	Min. 3x3 Stan. 15x15

Table 7.6: Characteristic values of all 3 stacks generations

7.4.4 FUJI 3D stack results

7.4.4.1 FUJI 3D stack at room temperature

The first stacks to be processed after the logic on DRAM test case were the FUJI based stacks. A FUJI die thinned down to 50 μm was stacked on a full thickness 700 μm die, back-to-face, as shown in figure 7.38. The FUJI stack used the same underfill as in the logic on DRAM test case, with slighlty different microbump dimensions. The upper Cu pad remained 30 μm in diameter, the Cu-Sn region was smaller equaling 15 μm in diameter and the bottom Cu pad spanned 25 μm in diameter. While the rest of the main stack characteristics, table 7.6, remained the same, the thicker top Si represented the main change. TSVs connecting top and bottom die were present, however the measured sensors on the bottom die were not connected. Therefore, only the top die FEOL was assessed.

Figure 7.39 presents a part of a layout on the top Si die that was used to assess the underfill-microbump mechanism on FUJI stacks. A 5x5 p-FET array with transistor width to length ratio 9 μm/8 μm was placed next to an underlying standalone microump. This array was labeled the microbump array. The p-FET array was placed 14 μm from the microbump center, therefore the underlying microbump projection just touched the beginning of the p-FET array zone. A 5 μm in diameter TSV was also processed next to the microbump array at 10 μm distance. In electrical measurements on the same p-FET arrays with no microbumps it was revealed that the TSV has only minor imapct on the first transistor in the middle row of the array, closest to the TSV. Another array was monitored with no microbumps or TSVs in its immediate surrounding, labeled remote p-FET array. A 4x2 array of microbump was present at a distance of 130 μm.

Electrical measurements of the microbump p-FET array and remote p-FET array before and after stacking are shown in figure 7.40. The array current shifts before stacking are normalized to the median value of their transistor array drain currents. The current shifts after stacking are normalized to the current value of the same transistor

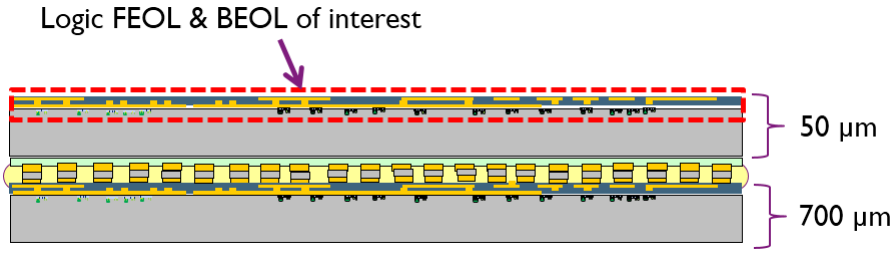


Figure 7.38: Second generation of stacks - FUJI test chip based stacks. 50 μm on 50 μm, back-to-face.

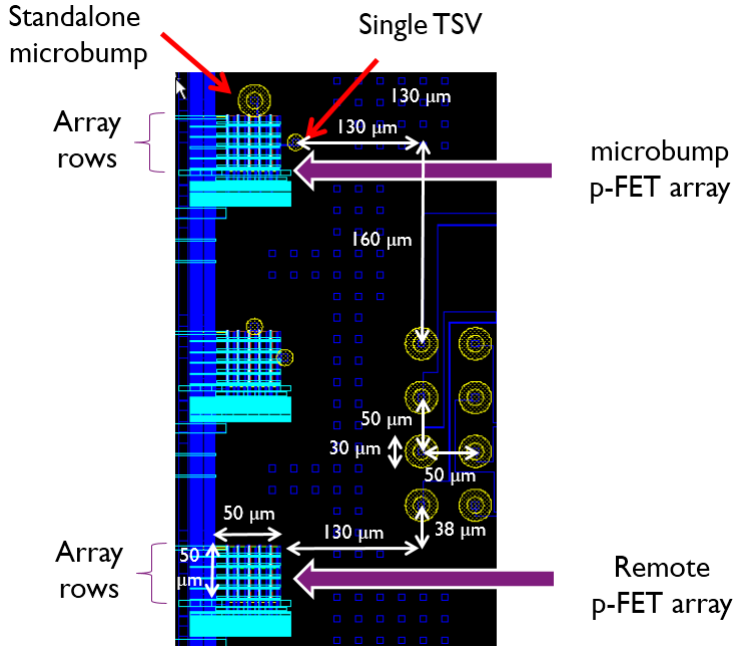


Figure 7.39: Part of a layout on the top Si die that was used to assess the underfill-microbump mechanism on FUJI stacks

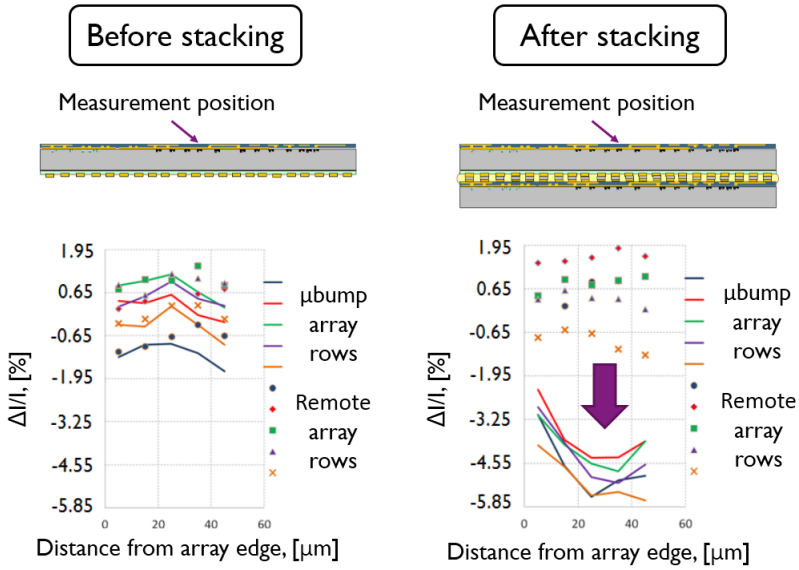


Figure 7.40: Electrical measurements of the microbump p-FET array and remote p-FET array before and after stacking are shown in figure 7.40. The array current shifts before stacking are normalized to the median value of their transistor array drain currents. The current shifts after stacking are normalized to the current value of the same transistor before stacking.

before stacking. A variation of current of approximately +/- 1% is seen before stacking in both arrays. Currents within one row of both arrays are approximately of same value.

After stacking a current shift of around -6% is visible, slightly changing between rows. Therefore, the stacking process through the underfill-microbump mechanism exerted a stress that caused a 6% decrease of p-FET transistor currents. p-FET transistors with current flow in [110] direction contain piezocoefficients which are of similar absolute value, but opposite signed. If the p-FET array was placed directly above the microbump where the in-plane stresses are of same magnitude and sign, minimal current shifts would be expected because the contributions of the two in-plane stress components to the p-FET current shift would cancel each other. In general, p-FETs are not desirable sensors for underfill-microbump mechanism monitoring. However, the p-FET array in this case is placed next to an underlying microbump, not directly above. Further away from the microbump center, one of the two in-plane stress components becomes dominant causing strongest current shifts in orbital type shapes, as seen in figure 7.9 for p-type Si.

Figure 7.41 compares the results of current shifts from the p-FET array obtained by finite element modeling and electrical measurements. The same material properties were used as in table 7.3. The position of the microbump and microbump array are highlighted on the simulated Si current shift pattern. The microbump array is positioned in the negative current shift orbital. The finite element model calculated stress

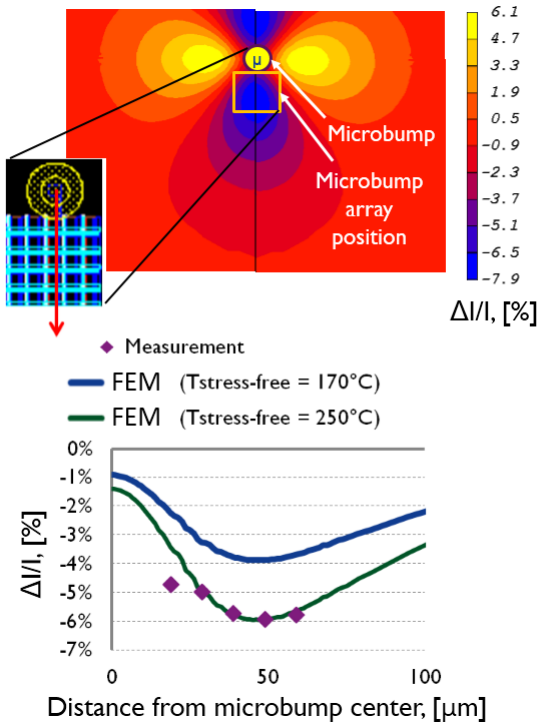


Figure 7.41: Results of current shifts from the central column of the microbump p-FET array obtained by finite element modeling and electrical measurements. The same material properties were used as in table 7.3.

values on the top Si surface which were transfered within the same model using the corresponding FUJI p-FET long channel piezocoefficients in table 5.1, obtained during in-plane stress calibration. In the graph in figure 7.41, the current shifts of the central column of the p-FET array were plotted, obtained with electrical measurements and by finite element modeling using equivalent stress free temperatures of 170ºC and 250ºC. The equivalent stress free temperature is the temperature within the simulation from which cooling of the modeled structure is commenced. Using an equivalent stress free temperature of 170ºC failed to match the electrically obtained current shifts, while usage of an equivalent stress free temperature of 250ºC resulted in a good match with electrically obtained current shifts. The equivalent stress free temperature of 250ºC corresponds to the stack bonding temperature. The results point in the direction that there is stress build-up present above the underfill glass transition temperature as well.

Figure 7.42 analyzes the remote p-FET array. Although there is no microbump present in its immediate surrounding, small current shifts are observed in the remote p-FET array as well. Finite element modeling revealed that the 4x2 microbump array visible in figure 7.39 to the right of the remote array is sufficiently close to the remote array to cause current shift. The position of the 4x2 microbump array and remote p-FET

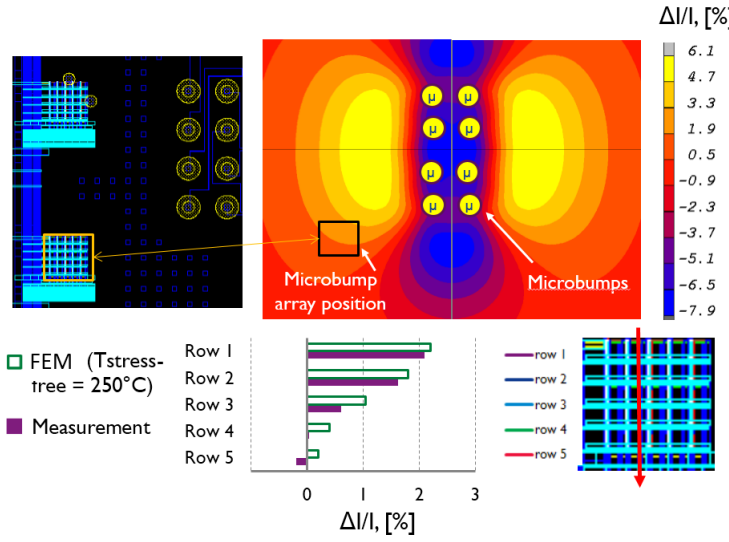


Figure 7.42: Analysis of current shifts observed on the remote p-FET array. Finite element simulations reveal that the 4x2 microbump array visible in figure 7.39 to the right of the remote array is in sufficient distance of the remote array to cause current shift. The position of the microbump array is indicated relative to the 4x2 microbump array.

array are highlighted on the simulated Si current shift pattern. The remote p-FET array is placed in the edges of the region of the positive current shift orbital. Furthermore, in figure 7.42 finite element and electrically obtained current shift values of the second column of the p-FET array are compared. The transistor in row 1, closest to the positive current shift orbital exhibits the highest current shift while the furthest transitor from the orbital exhibits the lowest current shift. Transistors in row 4 and 5 show a mismatch with the finite element model, possibly due to local design features surrounding these transistors.

The distance from the center of the closest microbump in the 4x2 array to the edge of the remote p-FET array is 130 μm, as seen in figure 7.39. According to figure 7.4 b), a 30 μm in diameter microbump exhibited an impact zone in Si of 140 μm. If the FUJI 30μm/15μm/25μm microbump is approximated as equivalent to the ETNA microbump, a part of the remote p-FET array with a distance of 130 μm from the 4x2 microbumps indeed lies within its impact range. The p-FET microbump array edge is at a distance of 206 μm from the first microbump in the 4x2 microbump array and is therefore out of the impact zone of those microbumps.

Direct stress extraction from the configuration of sensors on the FUJI die is not possible. The FUJI die consists of n-type and p-type devices oriented only in [110] direction. The p-FET microbump array as well consists of only [110] oriented transistors. Since out-of-plane stress on the top Si side is neglected, as discussed in section 7.1 and deducted from figure 7.4, the following minimal requirements need to be satisfied for

[001] surface Si in order to extract in-plane stress on the top Si side:

- Presence of two MOSFETS, one in longitudinal direction, the other in transverse direction
- a MOSFET array or transistor oriented in only one direction need to be placed directly above the underlying microbump which is either a standalone microbump or a microbump within a symmetric microbump array so that both in-plane stress components can be approximated as equal. This procedure can only be considered valid direclty above the underlying microbump close to the microbump center.

The transistors in the p-FET array in this case are oriented only in the [110] direction. Furthermore, the array is placed next to the underlying microbump, not above it. A direct extraction cannot be made. However, in order to get an idea of the magnitude of the stress, the following approximation can be made. Since the p-FET array is placed in the negative current shift orbital where one in-plane stress is dominant, the other in-plane stress can in first approximation be neglected. Using eq. 5.7 and neglecting one in-plane component and inputting a current shift of -6%, the stress results to 150 MPa. Although an approximate method, the value of 150 MPa corresponds fairly close to the 4.5 times stress drop with changing Si thickness from 25 μm to 50 μm, projected in figure 7.24. The exact ratio of the stress observed in the ETNA logic on DRAM case with a 25 μm thick die where 827 MPa was simulated and the approximated value from FUJI stacks where 150 MPa is estimated is 5.5. The approximated value of 150 MPa gives a fairly good sense of how the stress significantly dropped in the FUJI stack with primarily just changing the Si thickness from 25 μm to 50 μm compared to the previous logic on DRAM case.

7.4.4.2 FUJI 3D stack at elevated temperatures

A n-FET microbump array of the same configuration as the p-FET microbump array, with 8 μm x 9 μm transistors was monitored to assess underfill-microbump stress build-up at elevated temperatures. The n-FET array was placed in the same surrounding as the p-FET microbump array, next to a 30 μm/ 15 μm/ 25 μm in diameter microbump and a 5 μm in diameter TSV. The TSV, as with the p-FET array, had minor impact only on the closest transistor in the middle array row. The microbump was placed at a distance of 14 μm from the n-FET array edge. The electrical current of transistors in the n-FET microbump array was monitored at 25ºC, 55ºC and 85ºC. Figure 7.43 shows the distributions of current shifts of the n-FET microbump array at 25ºC and the layout of the n-FET array and its surrounding. Since the p-FET array from section 7.4.4.1 was used for stress monitoring at room temperature, in order to obtain the appropriate current shifts all of the transistors were referenced to their own values before stacking. When raising the temperature, the transistors drain current is impacted by the change in stress and the direct impact of temperature change. In order to separate these effects, assuming each transistor in the array has the same sensitivity to temperature, the drain currents of the whole array were referenced to the minimum drain current value within the array. This does not give the most precise method for tracking stress but it does

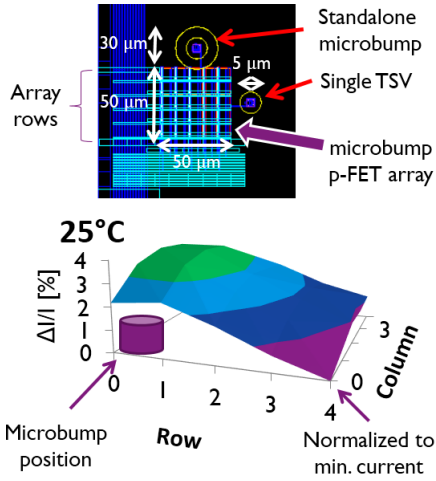


Figure 7.43: Current shifts of the n-FET microbump array at 25ºC next to the layout of the n-FET array and its surrounding. Currents are referenced to the minimum value within the array.

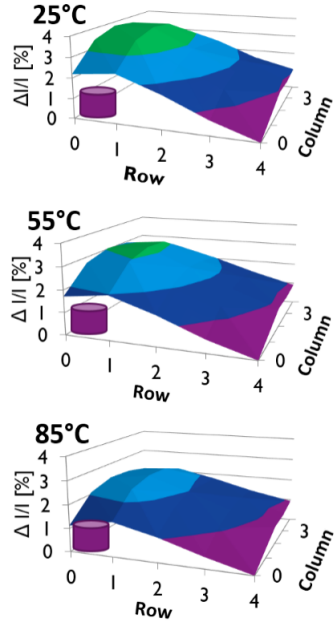


Figure 7.44: Current shift measurements on 25ºC, 55ºC and 85ºC all referenced to the minimum current value within the array. A clear drop in current shift with temperature due to stress relaxation is visible.

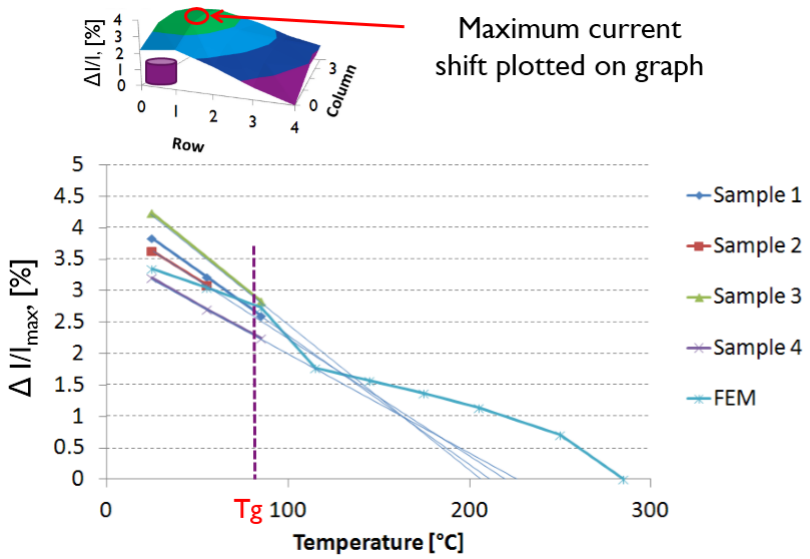


Figure 7.45: Maximum current shifts of the n-FET microbump arrays from 4 stacks versus temperature and current shifts obtained by the finite element model. The glass transition temperature of UF 6-A was 81 ºC and is highlighted in the graph. The maximum temperature at which the currents were electrically measured was 85 ºC due to limitations of the measurement equipment, which is in essence up to the glass transition temperature of the underfill.

ensure that at elevated temperatures, when normalizing to the minimum current, the current shifts are related only to the change in stress.

Figure 7.44 shows current shift measurements at 25ºC, 55ºC and 85ºC all referenced to the minimum current value within the array. A clear drop in current shift with temperature due to stress relaxation is visible. The current shift shape lowers with the maximum values dropping from 3.8% at 25ºC to 3.2% at 55ºC and 2.6% at 85ºC.

In a further study, n-FET microbump arrays from 4 FUJI stacks were measured at the same temperatures - 25ºC, 55ºC and 85ºC. In parallel, finite element simulations were performed to produce current shifts at the same n-FET array locations. The piezocoefficients for FUJI long channel n-FET transistors with channel width to length ratio of 8 μm/ 9 μm obtained through in-plane stress calibration were used to transfer in-plane stress to current shift.

Figure 7.45 plots the maximum current shifts of the n-FET microbump arrays from 4 stacks versus temperature and includes the current shifts obtained by the finite element model. The glass transition temperature of UF 6-A was 81 ºC and is highlighted in the graph. The maximum temperature at which the currents were electrically measured was 85 ºC due to limitations of the measurement equipment, which is in essence up to the glass transition temperature of the underfill.

All 4 stacks exhibited between 3 and 4% current shift at room temperature dropping

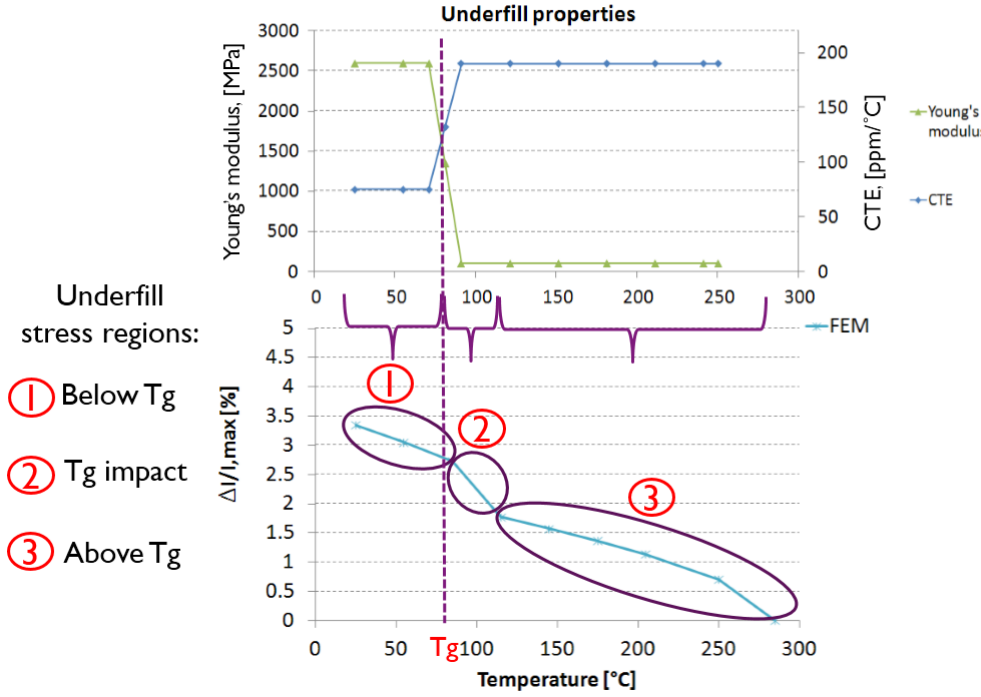


Figure 7.46: Comparison of the used underfill's Young's modulus and CTE with the generated current shift levels. Three regions of stress generation can be isolated.

linearly by about 1% at the highest measured temperature of 85ºC due to relaxation of stress at elevated temperatures. The glass transition temperature of the underfill was 81ºC, exactly at the limit of the electrical measurement setup which means direct monitoring above Tg was not possible. The linearly projected equivalent zero stress temperature according to electrical measurements was estimated to just above 200ºC. However, bonding is done at 250ºC and the underfill Young's modulus and CTE exhibit non-linear changes crossing the Tg which is why linear stress build up above Tg does not seem realistic. The finite element model considered cooling of the stack from a temperature above 250ºC in order to take into account the cure shrinkage of the underfill before actual structure cooling starts.

According to the finite element model, from bonding temperature the stress builds up slightly non-linearly to 1.5% current shift reaching the zone around Tg. Just before Tg, the changing underfill properties seem to cause a steeper stress increase in Si and below Tg the stress linearly reaches a value above 3% current shift, which is within the range of the electrical measurements. The electrical measurements from figure 7.45 clearly show that not all stress has been generated between the underfill's T_g and room temperature. Stress is being generated above the underfill's T_g as well. The stress build up above T_g contributed to more than 50 % of the total stress. The finite element model current shifts at room temperature match with the experimentally obtained

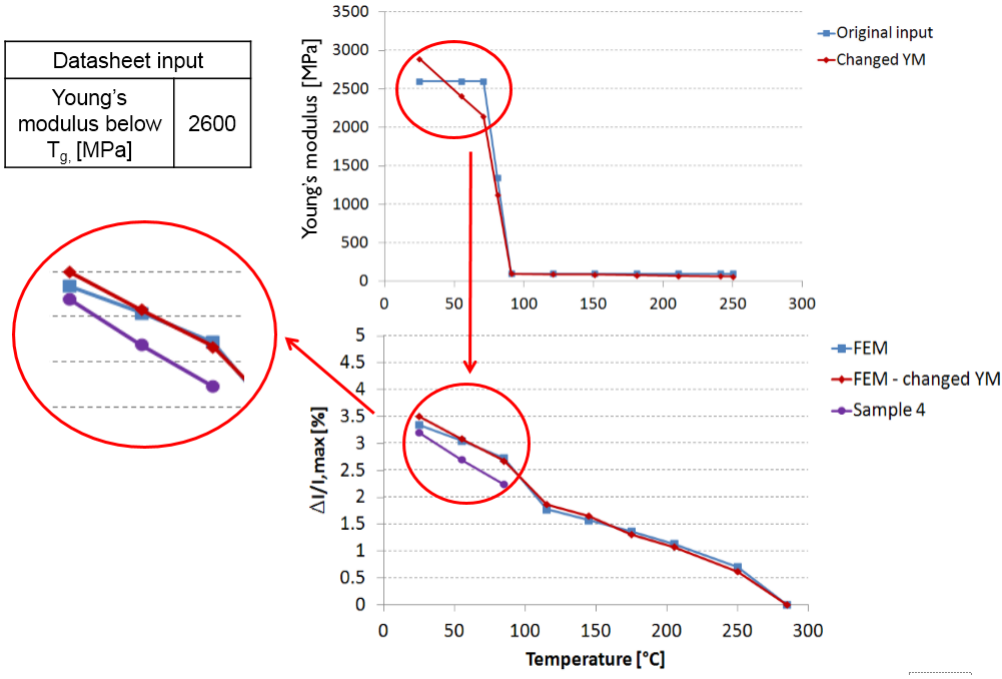


Figure 7.47: An exercize where the underfill's Young's modulus was altered to fit the slope below Tg of the current shift obtained through electrical measurements. The model accuracy is highly dependent on the accuracy of its input parameters.

current shifts at room temperature. Furthermore, at higher temperatures above T_g, the finite element model, although limited in accuracy due to the complexity and accuracy of inputted material properties, also indicates significant stress build-up. Therefore, the underfill-microbump stress build up above T_g should not be neglected.

Figure 7.46 compares the used underfill's Young's modulus and CTE with the generated current shift levels. Three regions can be isolated in comparison with the changing Young's modulus and CTE properties:

- Stress build-up above T_g, region 3
- Stress build-up around T_g, region 2
- Stress build-up below T_g, region 1

The stress build-up around and above the underfill Tg is sensitive to the underfill's Young's modulus and CTE changes. The exact Young's modulus and CTE are hard to know as they are not always available from DMA and TMA curves, but only publicly available datasheets, if not characterized in house. A slight change in Young's modulus or CTE changes the slope of the linear stress build up as well which can then better be matched to the slope of the electrical measurements. Figure 7.47 presents results

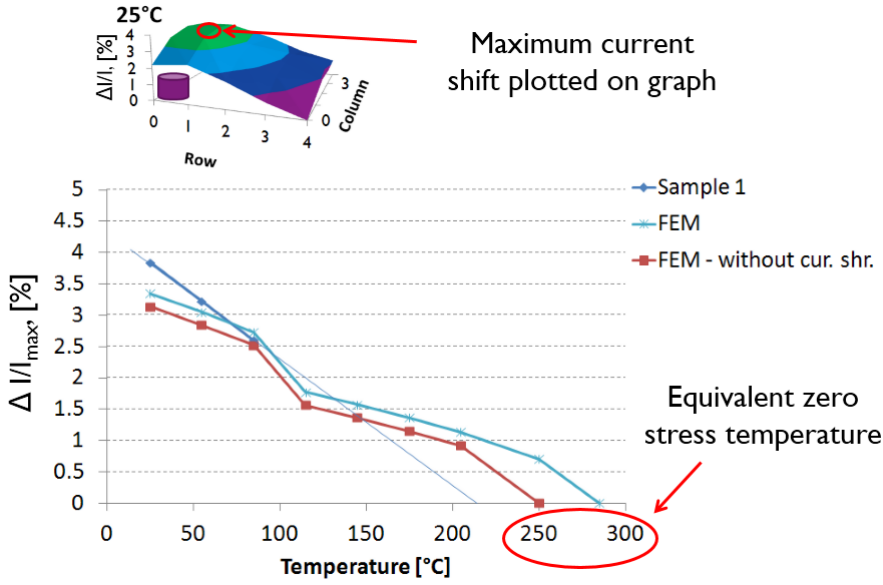


Figure 7.48: Simulated current shifts with and without included underfill cure shrinkage versus electrical measurements on one sample. Inclusion of cure shrinkage gives rise to current shift. The impact of cure shrinkage should not be neglected.

of an exercise where the underfill's Young's modulus was altered to fit the slope of the current shift below T_g obtained through electrical measurements. After alteration of the underfill's Young's modulus from a constant value of 2600 MPa below T_g to a linearly rising value from 2200 MPa to 2900 MPa, the slope of the finite element obtained current shift matched the slope of the electrically obtained current shifts. Young's modulus values obtained from DMA curves are not always constant below Tg and with certain underfills a certain slope below Tg can be expected. DMA and TMA curves for most underfills in figure 7.15 were obtained as input for finite element simulations. Unfortunately, data for UF 6-A was available only from a datasheet. For this reason, obtaining precise underfill material properties is of high importance.

Figure 7.48 serves as a background for a discussion related to the underfill's cure shrinkage. All 3D IC stacks were in fact modeled with an effect from cure shrinkage, according to eq. 7.3. Although a simplistic inclusion of cure shrinkage, models with included cure shrinkage matched closer the electrical measurements. In the case of FUJI stacks, finite element obtained current shift dropped by 0.3%. The difference varies with different used underfills. As visible from figure 7.48, in order to include underfill cure shrinkage, the equivalent zero stress temperature used on the finite element model from which the 3D stack model was cooled was raised according eq. 7.3. More substantial evidence on the impact of underfill's cure shrinkage was not gathered and it can be argued that a more complex underfill model without the usage of cure shrinkage will also result in better matching with electrical measurements. However, until proven

otherwise, cure shrinkage in terms of stress build-up should not be neglected.

7.4.5 PTCQ 3D stacks results

The FUJI stacks from section 7.4.4 were followed by stacks based on the stress dedicated test chip PTCQ. The design of the PTCQ die is presented in section 7.4.2 and the characteristics of the PTCQ 3D stacks summarized and compared to the ETNA and FUJI stacks in table 7.6. The PTCQ stack maintains the top die thickness at 50 μm as the FUJI stack but lowers the microbump pitch to 40 μm. Furthermore, a new underfill is used for the PTCQ stacks, UF 9-C, which however induces almost the same amount of stress as the previous underfill, UF 6-A, as visible from figure 7.15. UF 9-C was chosen by the stacking development team primarily due to compatibility with the in-house stacking process, rather than its thermo-mechanical behavior.

PTCQ n-type and p-type transistors present in all cells and used as global and local transistors were calibrated to in-plane stress using 4-point bending. The results are summarized in figure 7.49 and table 7.7. Out-of-plane calibration on the same PTCQ n-type transistors was presented in section 5.2.2. The wafer orientation of the PTCQ die is shifted by 45º compared to all previous test chips. Transistors in longitudinal direction now have current flow in [100] and transverse transistors in [010], as visible from the illustration above table 7.7. Table 7.7 also includes values from Smith [93] for [100] and [010] directions. While the stress calibration results from previous test vehicles showed either resemblence to the Smith piezocoefficients or a trend was present that could be related to the Smith coefficients, some PTCQ in-plane piezocoefficients are considerably different. For the n-type and p-type transistor, stress calibration in transverse direction to the current revealed piezocoefficients of opposite signs to Smith's piezocoefficients. The n-type longitudinal calibration exhibited a piezocoefficient value 3 times lower than the Smith's piezocoefficient equivalent while the p-type longitudinal piezocoefficient was very close to Smith's piezocoefficient. While lower sensitivities compared to Smith's piezocoefficients have been previously observed in section 5, the reason for transverse n-type and p-type piezocoefficients with opposite signs compared to Smith's piezocoefficients is to this point unknown.

When out-of-plane stress is neglected, eq. 5.7 and eq. 5.8 can be rewritten in terms of individual stress components using experimentally obtained piezocoefficients from figure 7.49 and table 7.7. Eq. 7.4 and 7.5 extract longitudinal stress in [100], σ_l, and transverse stress in [010], σ_t, from a set of longitudinal and transverse n-type transistors with channels oriented in [100] and [010] directions and currents referred to as $I_{n,0}$ and $I_{n,90}$, respectively. Eq. 7.6 and 7.7 extract the same components of stress based on the same configuration with p-type transistors.

$$\sigma_l = 54.6 \cdot \frac{\Delta I_{n,0}}{I_{n,0}} - 33.9 \cdot \frac{\Delta I_{n,90}}{I_{n,90}} \tag{7.4}$$

$$\sigma_t = -33.9 \cdot \frac{\Delta I_{n,0}}{I_{n,0}} + 54.6 \cdot \frac{\Delta I_{n,90}}{I_{n,90}} \tag{7.5}$$

$$\sigma_l = 14.8 \cdot \frac{\Delta I_{p,0}}{I_{p,0}} - 52.3 \cdot \frac{\Delta I_{p,90}}{I_{p,90}} \tag{7.6}$$

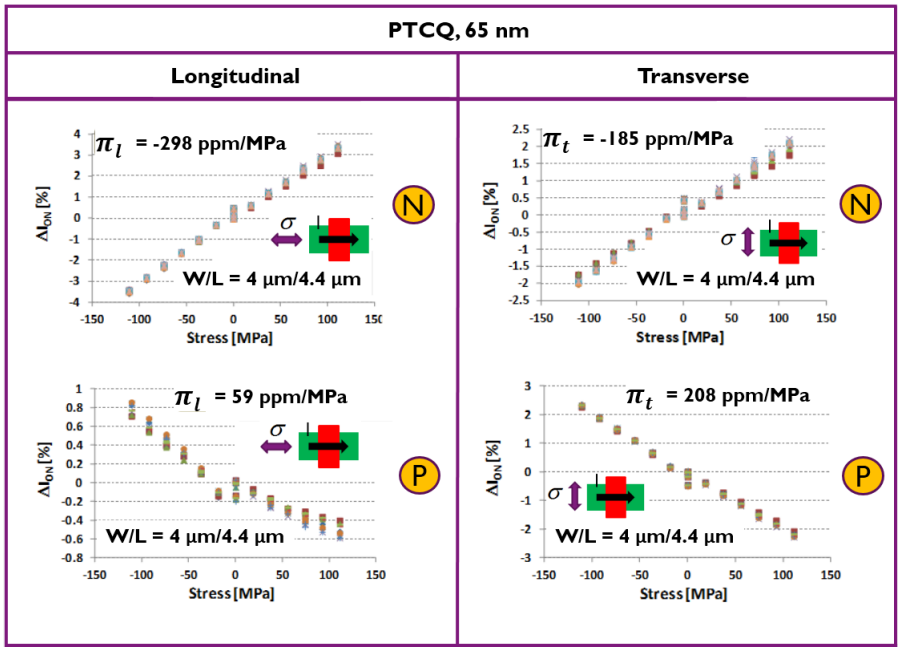
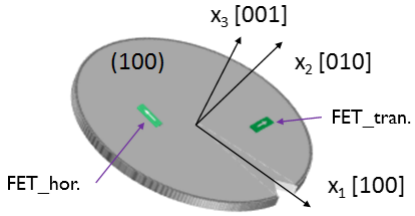


Figure 7.49: Current shift calibration results to in-plane stress for n-type and p-type MOSFETs evaluated on the PTCQ test chip



Technology node	FET type	Width/Length (nm/nm)	Piezocoeff. type	Piezocoeff. Saturation (ppm/MPa)	Piezocoeff. Linear (ppm/MPa)	Smith piezocoeff. (ppm/MPa)
65 nm (PTCQ)	N, long	500/900	longitudinal	-298	-	-1022
	N, long	500/900	transverse	-185	-	534
	P, long	500/900	longitudinal	59	-	66
	P, long	500/900	transverse	208	-	-11

Table 7.7: In-plane piezocoefficients of the PTCQ n-type and p-type MOSFETs, corresponding to figure 7.49

$$\sigma_t = -52.3 \cdot \frac{\Delta I_{p,0}}{I_{p,0}} + 14.8 \cdot \frac{\Delta I_{p,90}}{I_{p,90}} \tag{7.7}$$

In eq. 7.4 -7.7 $\frac{\Delta I_{n,0}}{I_{n,0}}$ stands for n-type transistor current shift in longitudinal direction [100], $\frac{\Delta I_{n,90}}{I_{n,90}}$ stands for n-type transistor current shift in transverse direction [010], $\frac{\Delta I_{p,0}}{I_{p,0}}$ stands for p-type transistor current shift in longitudinal direction [100] and $\frac{\Delta I_{p,90}}{I_{p,90}}$ stands for p-type transistor current shift in transverse direction [010].

Eq. 7.4 -7.7 are useful for stress extraction in Si in biaxial stress areas with negligible out-of-plane stress such as on the Si surface opposite of the microbump or in a 2D package where the FEOL is facing the mold compound side. These equations can also be applied to the biaxial test discussed in section 5.4. In a state where both in-plane stresses and out-of-plane stress is present but where two in-plane stresses are of similar value, eq. 5.5 and eq. 5.6 incorporating all normal stresses can be again reduced to two-stress equations. Eq. 5.5 and 5.6 can be rewritten as

$$\frac{\Delta I_1}{I_1} = (\pi_l + \pi_t)\,\sigma_i + \pi_v \sigma_v \tag{7.8}$$

$$\frac{\Delta I_2}{I_2} = (\pi_l + \pi_t)\,\sigma_i + \pi_v \sigma_v \tag{7.9}$$

where σ_i represents the in-plane stress. Eq. 7.8 and 7.9 are essentially equivalent equations meaning a set of one longitudinal and one transverse transisor, n-type

or p-type do not contribute in extraction of in-plane stress σ_i and out-of-plane stress σ_v. In order to extract in-plane stress σ_i and out-of-plane stress σ_v a set of one n-type and one p-type transistor can be considered. In this instance, from an arithmetic viewpoint, it makes no difference if one or the other transistor is in longitudinal or transverse direction. A combination of longitudinal-longitudinal, longitudinal-transverse or transverse-transverse transistors can be chosen, as long as one transistor is n-type and the other is p-type. Eq. 7.10 and 7.11 present a set of equations with one transverse n-type transistor, $I_{n,90}$, and one transverse p-type transistor, $I_{p,90}$, respectively. This equation system represents two equations with two unknowns from which the in-plane stress σ_i and out-of-plane stress σ_v can be extracted.

$$\frac{\Delta I_{n,90}}{I_{n,90}} = \left(\pi_{l,n} + \pi_{t,n}\right)\sigma_i + \pi_{v,n}\sigma_v \tag{7.10}$$

$$\frac{\Delta I_{p,90}}{I_{p,90}} = \left(\pi_{l,p} + \pi_{t,p}\right)\sigma_i + \pi_{v,p}\sigma_v \tag{7.11}$$

7.4.5.1 Global stress

PTCQ long channel n-FETs and p-FETs from cell 1 and cell 2 were measured to analyze global stress occuring across the die after 3D IC stacking. The current shifts were calculated by comparison to wafer level measurements. N-type and p-type transitors from cell 1 and 2 were placed in between microbumps as shown in figure 7.50 a). The underfill is present in the actual 3D stack but has been removed from figure 7.50 for better visibility. In a three dimensional perspective, the transistors were placed within a microbump array with equal distance of 28 μm to all 4 surrounding microbumps. In this area, the two in-plane stresses are considered to be equal. Therefore, eq. 7.10 and 7.11 can be used to extract the in-plane and out-of-plane stress in the position of the global sensors across the die. Figure 7.50 b) presents extracted in-plane and out-of-plane stress values on both tiers. All stresses on tier 1 and tier 2 are below 100 MPa. In-plane stresses on tier 1 and tier 2 vary across the die between 40 MPa and 100 MPa. The out-of-plane stress seems to be slightly lower than the in-plane stresses, however within the range of in-plane stress measurements. The out-of-plane stress on tier 2 varies around zero, confirming out-of-plane stress on the opposite side of the microbump is negligible. No global trend can be recognized. Stress variations along the die can be attributed to local structures impacting the sensors.

7.4.5.2 Local stress

PTCQ long channel n-FETs from cell 3 were measured to analyze local stress from the underfill-microbump mechanism and the Cu pad after 3D IC stacking. Figure 7.51 presents the layout of the 7x8 n-FET array and individual surrounding n-FETs in the surrounding of a) microbumps and b) a single Cu pad. The n-FET array is approximately 40 μm x 40 μm with transistor width to length being 4 μm x 4.4 μm. The Cu pad which is part of the later created microbump on the n-FET array side spans 15 μm. In the second case, the Cu pad is placed above the n-FET array and spans 50 μm in diamater and 5 μm in height. The microbumps in the surrounding are placed

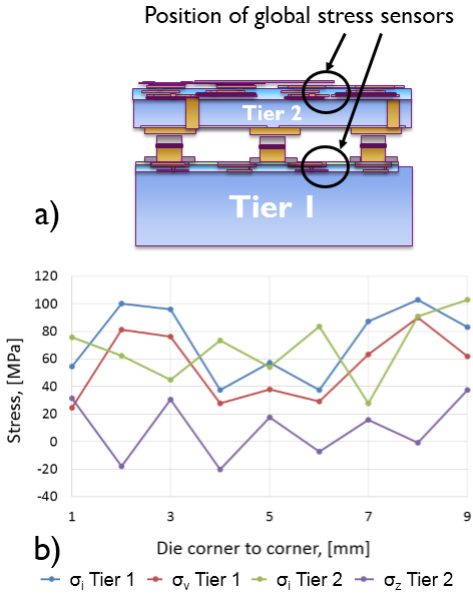


Figure 7.50: Extraction of in-plane and out-of-plane stress on the PTCQ stack: a) position of n-type and p-type global sensors and b) extracted stress values on tier 1 and tier 2

at 40 μm pitch from each other. The microbump array is large reaching at least 15x15 microbumps, except in limited number of cells on the very edge of the die.

Figure 7.52 summarizes the positions of local n-FET arrays and the structural combinations they have been applied to:

- in the presence of a microbump - directly below the microbump and on the opposite side of the microbump
- in the presence of a Cu pad, no microbump present - on tier 1 and directly below the Cu pad on tier 2
- in the presence of both microbump and Cu pad - directly below the microbump and on the opposite side of the microbump below a Cu pad

The underfill is present in the actual 3D stack but has been removed from figure 7.52 for better visibility. Figure 7.53 summarizes the measured n-FET array current shifts from the combinations in figure 7.52. Current shifts from 8 transistors of the 4th column are plotted, as indicated by the red arrow in figure 7.51 a) and b). The reference values are taken before stacking on wafer level.

Figure 7.53 a) shows the response of the n-FETs below the microbump and directly above the microbump, case a) in figure 7.52. On tier 2, on the opposite side of the microbump, around 6% current shift is detected. On the opposite side of the microbump in-plane stress is dominant. In case of ETNA stacks, clear current shift curvature was

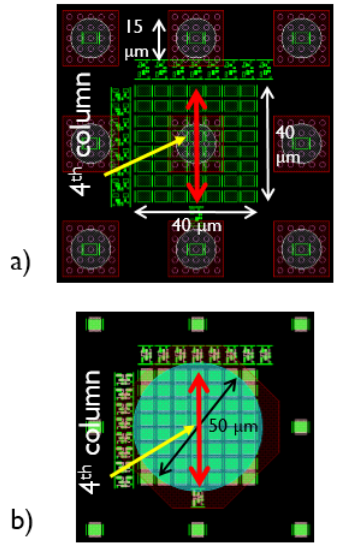


Figure 7.51: 7x8 n-FET array and individual surrounding FETs acting as local stress sensors on the PTCQ die: a) the n-FET array in a surrounding of microbumps and b) the n-FET array below a single Cu pad.

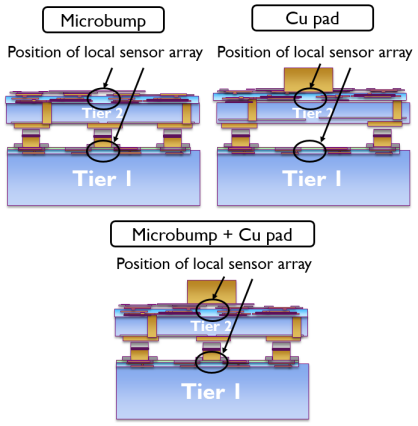


Figure 7.52: Positions of local n-FET arrays and the structural combinations they have been applied to: a) array on tier 1 below a microbump and on tier 2 on the opposite side of the microbump, b) array on tier 1 with no microbump or Cu pad and on tier 2 below a Cu pad and c) array below a microbump on tier 1 and on the opposite side of the microbump below a Cu pad

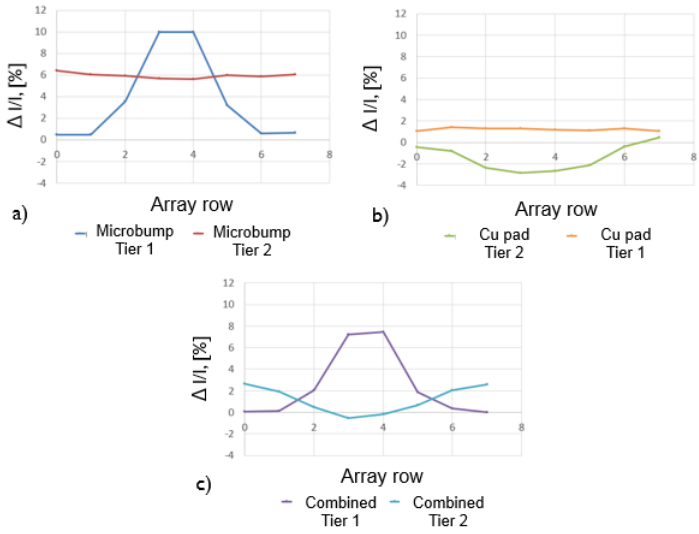


Figure 7.53: The measured n-FET array current shifts from the combinations in figure 7.52. Current shifts from 8 transistors of the 4th column are plotted, as indicated by the red arrow in figure 7.51 a) and b). The current reference values are taken before stacking on wafer level.

visible on the opposite side of the microbump with highest current directly above the microbump position. In this PTCQ case, the die is twice as thicker, at 50 μm, the stress is significantly lower and the pronounced current shift shape dissapears. Instead, the current shift remains steady across the whole array. If the two in-plane stress components are considered equal, using equation 5.7 for a single transistor when out-of-plane stress is neglected and experimentally obtained PTCQ piezocoefficients for n-FETs from table 7.7, the extracted in-plane stress in Si on the opposite side of the microbump equals to 125 MPa. This value is in range of the stress extracted on the FUJI stacks. The FUJI stacks with same top Si thickness and an underfill which induces similar amounts of stress exhibited a stress of 150MPa. The small microbump pitch might have had an impact on the slightly lower stress in the PTCQ stack.

Measurements on tier 1 for the first time reveal current shifts directly below the microbump. A current shift curvature is present with values up to +10%. As previously discussed during transistor out-of-plane stress calibration in section 5.2.2 and figure 5.10, directly below the microbump compressive out-of-plane stress becomes dominant. Therefore, the current shift is predominantly impacted by the out-of-plane stress. Figure 7.54 repeats the finite element simulation results from section 5.2.2, figure 5.10, plotting out-of-plane stress and in-plane stress over a path below 5 microbumps in tier 1 of the PTCQ stack. The results in figure 7.54 a) point to high compressive out-of-plane Si stress of 400 MPa below microbump positions and small tensile out-of-plane stress in between the microbumps of around 100 MPa. The in-plane stress results, 7.54 b), point to compressive in-plane Si stress below the microbump approximately 3.6 times

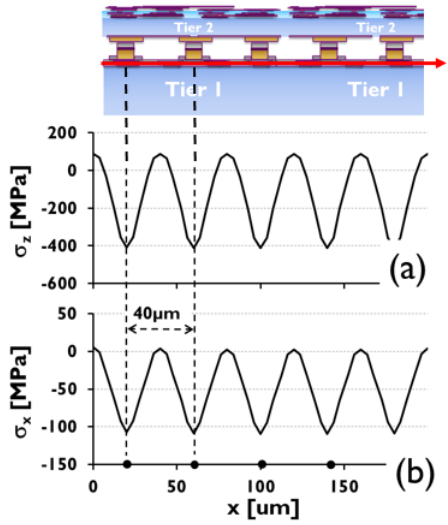


Figure 7.54: Simulated out-of-plane stress and in-plane stress over a path below 5 microbumps in the tier 1 of the PTCQ stack. The results in figure 7.54 a) point to high out-of-plane Si stress of 400 MPa below microbump positions and small out-of-plane stress in between the microbumps of around 100 MPa. The in-plane stress results, 7.54 b), point to compressive in-plane Si stress below the microbump approximately 3.6 times lower than the out-of-plane stress. The in-plane stress values between microbumps is negligible.

lower than the out-of-plane stress. The in-plane stress values between microbumps is negligible. In comparison with the results from tier 1 global sensors lying in the same position between the microbumps in figure 7.50, the finite element model underestimates the global stress. The global sensors extracted in-plane stress values between 40 MPa and 100 MPa, depending on local features in the die. In the simulation, Si was taken as isotropic and Cu plasticity was not included. Enhancements of the model, such as including Si as orthotropic and Cu plasticity could provide a more accurate view on absolute stress values. Regardless, the finite element model in figure 7.54 does aid in presenting the trends of dominant compressive out-of-plane stress below the microbump compared to the previously observed negligible out-of-plane stress and dominant in-plane stress in Si on the opposite side of the microbump.

Figure 7.55 compares out-of-plane stress values in Si directly below the microbump, as in figure 7.52 a), obtained with 2 different methods: the electrical measurements from figure 7.53 a) - tier 1 and finite element model from figure 7.54 a). At lower stresses the extracted stresses vary while the methods match at compressive stress values. The out-of-plane stress extraction from electrical measurements was based on the out-of-plane stress calibration from section 5.2.2 and piezocoefficient values in table 5.5. The initial piezocoefficient value of 300 ppm/MPa is taken for current shift to stress conversion as it incorporates presence of in-plane stress in its value. Direct out-

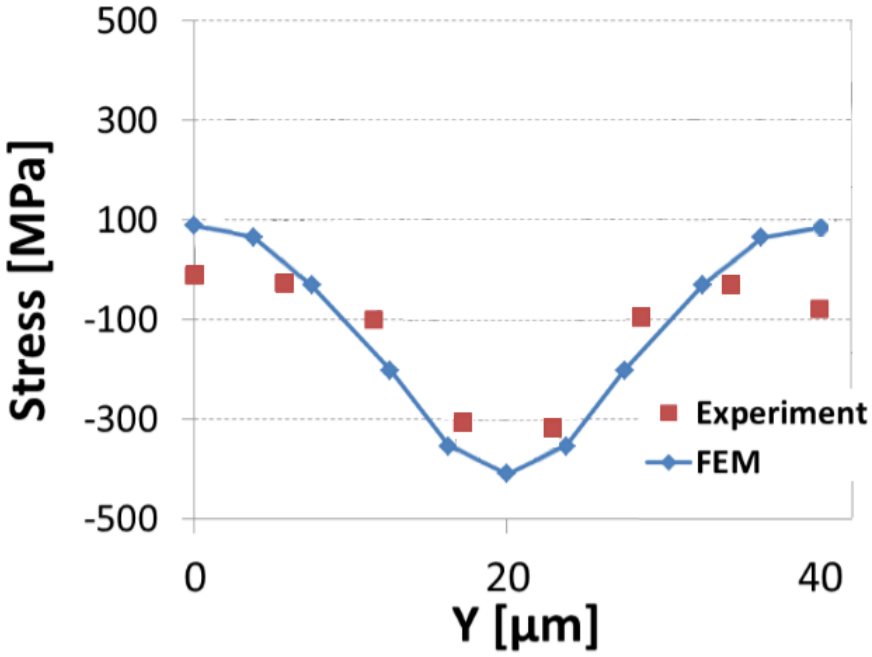


Figure 7.55: Out-of-plane stress values in Si directly below the microbump, as in figure 7.52 a), obtained with 3 different methods: the electrical measurements from figure 7.53 a) - tier 1, finite element model from figure 7.54 a) and micro-Raman measurements. At lower stresses the extracted stresses vary while all 3 methods match at compressive stress values higher than 300 MPa.

of-plane stress extraction is otherwise not possible as only one transistor type in one orientation is available.

Figure 7.53 b) presents results on tier 2 with a Cu pad and with no microbump present on tier 1. The current shift responses on tier 1 exhibit a slight positive shift of +1.5% while tier 2 current shifts exhibit a small curvature with negative current shift reaching -3%. The case on tier 1 is equivalent to the position modeled between the microbumps in figure 7.54 and to the position of extracted global stress between the microbumps. When stress extracted from the global stress sensors in figure 7.50 is used as input to equation 7.10, describing local n-FET array transistor response, the calculated current shifts of the n-FET array lie within range of the actual measured current shifts. If 50 MPa of in-plane stress and 20 MPa of out-of-plane stress is taken from the extracted stress values in figure 7.50, the calculated current shift is exactly +1.5%, as observed with measurements. The negative current shift below the Cu pad on tier 2 indicates presence of compressive stress.

Figure 7.53 c) presents current shift results directly below the microbump and opposite of the microbump in the presence of a Cu pad. The same current shift curvature on tier 1 is observed as without the Cu pad on the opposite side. On tier 2, the compressive impact of the Cu pad from 7.53 b) is impacted by the tensile in-plane stress from 7.53 a) resulting in a positive shift of the curve observed in figure 7.53 b) by about 2%.

Figure 7.56 analyses further the impact of the Cu pad. Figure 7.56 a) compares the current shifts on tier 2 with and without microbump presence and the forementioned positive shift of the current shift curve. Figure 7.56 b) presents the same on tier 1. While tier 1 is dominantly impacted by microbump presence, the resulting stress on tier 2 was still vague. An initial hypothesis was made that stresses below the Cu pad without the presence of the microbump is created by possible local bending by pushing on the Cu pad during stack bonding, illustrated on figure 7.56 next to the graphs. However the current shift curvature below the Cu pad, although shifted to positive values, is present even with an existing underlying microbump which would prevent the local bending. Therefore it is unlikely that local bending after stack bonding causes this stress. Furthermore, it is unlikely that cooling of the structure and shrinking of the Cu pad alone would cause sufficient stress to cause visible current shifts in Si.

In order to observe the evolution of current shift under a loaded Cu pad, n-FET array currents were measured during application of out-of-plane stress on the Cu pad. Out-of-plane stress was applied with a delaminator over a Si cube, as in section 5.2.2. Figure 7.57 summarizes the results of n-FET array current shift monitoring under a loaded Cu pad. Initially, at zero equivalent stress, reference current values from each transistor in the array are taken. During loading, the current shifts on the 4th column of the n-FET array are presented. With rising out-of-plane compressive stress, rising positive current shift is observed, consistent with the obtained piezocoefficient sign from section 5.2.2, table 5.5. After a maximum of 600 MPa and highest positive current shifts within the n-FET array, the stress is lowered to 300 MPa. At this stress, n-FET array current shifts vary around zero. Upon complete release of the load, negative current shifts of -2% remain corresponding to tier 2 results from 7.53 b) and 7.56 a). Therefore, irreversible deformation occurs which leaves stress below the Cu pad resulting in remaining negative current shfits on the local sensors beneath. The yield stress of Cu was previously extracted with nano-indentation on the Cu pad, resulting in

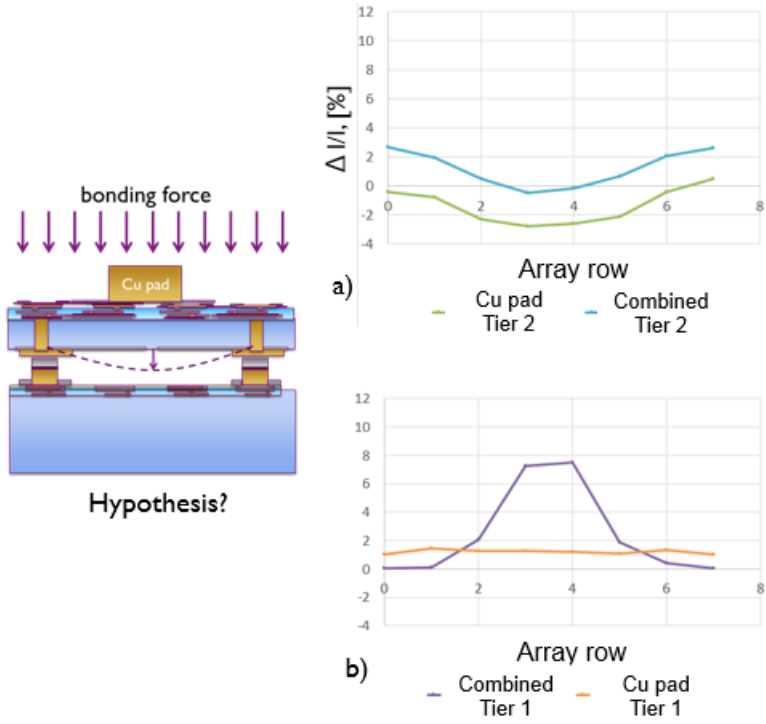


Figure 7.56: Analysis of the impact of the Cu pad. Figure 7.56 a) compares the current shifts on tier 2 with and without microbump presence. Figure 7.56 b) presents the same on tier 1. An initial hypothesis was made that stress below the Cu pad without the presence of the microbump is created by possible local bending by pushing on the Cu pad during stack bonding, illustrated next to the graphs.

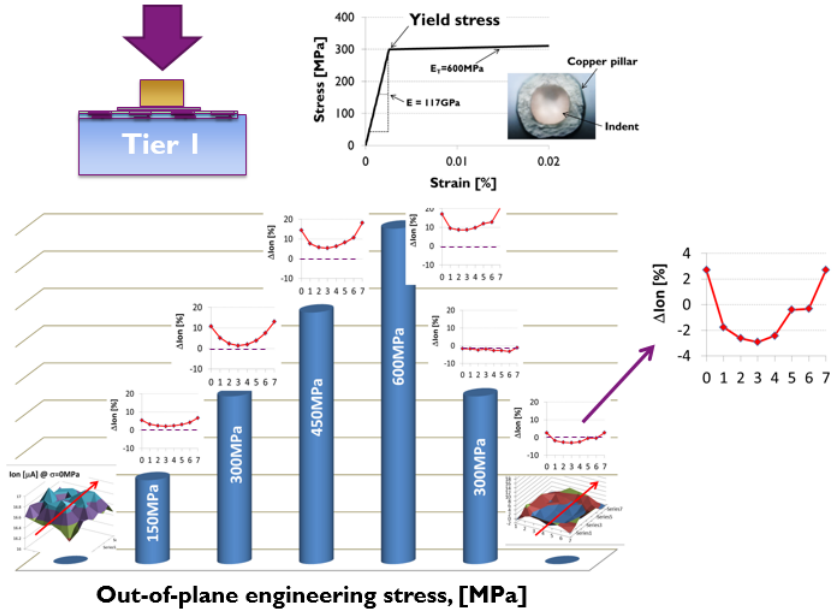


Figure 7.57: N-FET array current shift monitoring under a loaded Cu pad. Out-of-plane stress was applied with a delaminator over a Si cube, as in section 5.2.2. After removal of external out-of-plane stress, current shifts still remain due to plastic deformation of Cu.

3D IC stack	Top Si die thickness	Microbump pitch	Microbump array	Induced stress
ETNA	25 µm	> 200 µm	Standalone microbump	827 MPa
FUJI	50 µm	> 200 µm	Standalone microbump	150 Mpa
PTCQ	50 µm	40 µm	Min. 3x3 Stan. 15x15	125 MPa

Table 7.8: Main stack characteristics which were altered on the 3 test vehicles and lead to stress decrease in position illustrated in figure 7.58. An additional column is added with the extracted stress values. A significant decrease in stress is accomplished from 827 MPa on the ETNA stacks to 125 MPa on the PTCQ stacks.

300 MPa. Since the applied stress was higher than the yield stress of Cu, it is likely that the remaining current shifts are a result of plastic deformation of the Cu pad.A negative current shift according to eq. 5.5 and piezocoefficients from figure 7.49 and table 5.5 suggest a possibility of:

- acting dominant compressive in-plane stresses
- acting dominant tensile out-of-plane stress

Presence of tensile out-of-plane stress is not logical, as the out-of-plane stress was released and there is no common ground for presence of tensile out-of-plane stress after Cu pad deformation. The assumption is therefore that the negative current shifts are a result of dominant compressive in-plane stress below the Cu pad. Translated to the 3D stack, it is likely that during bonding plastic deformation of Cu pad occurs leading to the same stress generation. Compressive in-plane stress most likely occurs when, during out-of-plane compressive bonding loads, the bottom edges of the deformed Cu pad start pushing inwards in such a manner also compressing the BEOL and Si beneath.

7.4.6 Stress evolution - from ETNA to FUJI and PTCQ

The ETNA and FUJI stacks comprised sensors in the thin Si die on the opposite side of the microbump. PTCQ stacks enabled measurements on the opposite side of the microbump and directly below the microbump as well. It also enabled stress extraction between the microbumps and assessment of Cu pad impact. Evolution of Si stress in test vehicles ETNA, FUJI and PTCQ is presented by comparing stress levels from the only common stress extraction position in all test vehicles, on the opposite side of the microbump, as illustrated in figure 7.58. Table 7.8 states the main stack characteristics which were altered on the 3 test vehicles and lead to stress decrease in position illustrated in figure 7.58. An additional column is added to table 7.8 with the extracted stress values. A significant decrease in stress is accomplished from 827 MPa on the ETNA stacks to 125 MPa on the PTCQ stacks.

A large stress decrease is observed from ETNA stacks to FUJI stacks due to increase of the top thin die thickness from 25 μm to 50 μm. The stress from the FUJI

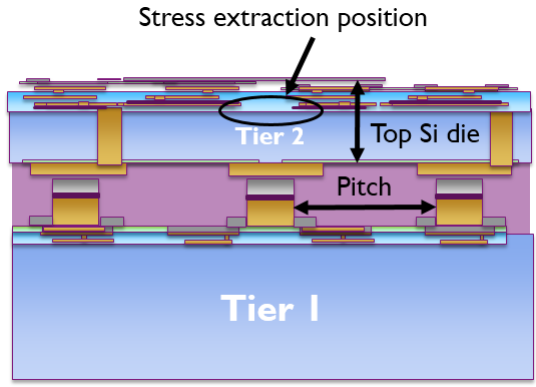


Figure 7.58: Evolution of Si stress in test vehicles ETNA, FUJI and PTCQ is presented by comparing stress levels from the only common stress extraction position in all test vehicles, on the opposite side of the microbump

stacks and PTCQ stacks are within range. The PTCQ stacks use an underfill which causes slighlty higher stress than in FUJI stacks, however the addition of lowered microbump pitch down to 40 μm and larger microbump arrays result in slightly lower stress than in FUJI stacks.

7.5 Summary

High current shifts above 40% were initially observed on n-FET transistors within a 3D IC stack, placed on the thin Si die, on the opposite side of the microbump. The initial stack consisted of a 25 μm thick die stacked on a 550 μm thick die. This current shift was linked to stress generated by the underfill-microbump stress mechanism. During cooling of a 3D IC stack after die bonding, the underfill as the material with the highest CTE in the surrounding, shrinks and pulls the thin Si die over the underlying microbumps causing local warpage of the thin Si die. The underfill microbump stress mechanism is a CTE mismatch driven mechanism initiated by the underfill material. The 40% current shift was linked to 827 MPa of Si stress.

In total three 3D IC stack generations were assembled: based on the ETNA test chip, the FUJI test chip and the PTCQ test chip. ETNA and FUJI are logic based test chips primarily processed for research on advanced FEOL devices, while PTCQ is a test chip dedicated to research on FEOL mechanical stress in 3D IC stacks and 3D IC packages. In the first 3D IC stack generation, ETNA based, and in the second 3D IC stack generation, FUJI based, stress in Si was analyzed on the thin die on the opposite side of the microbump. In the third 3D IC stack generation, PTCQ based, stress was analyzed in multiple positions, including the Si side opposite of the microbump, under the microbump and between microbumps. Additionaly, the stress impact of Cu pads used to create microbumps was assessed.

In summary, stress in Si through three 3D IC stack generations was assessed in the

following locations:

- on the opposite side of the microbump
- below a microbump
- between microbumps
- beneath a Cu pad
- combinations of a microbump and a Cu pad

On the opposite side of the microbump, tensile in-plane stress components in Si are dominant. Compressive out-of-plane stress becomes dominant in Si below the microbump. Generated stress patterns in Si from the underfill-microbump mechanism on the opposite side of the microbump are circular. Negligible amounts of stress are observed in Si if an underfill is not present.

N-type and p-type Si and in that sense n-type and p-type MOSFETs have a distinctively different response to mechanical stress. The circular stress in Si on the opposite side of the microbump generated by the underfill-microbump mechanism creates circular, dominantly positive current shift patterns for n-type Si and orbital current shift patterns for p-type Si with regions of negative and positive current shift. In finite element models, simulated stress values are linked to current shift using piezocoefficients obtained through stress calibration. During stress extraction using MOSFETs, using experimentally obtained piezocoefficients, measured current shift is transferred to in-plane and out-of-plane stress components.

Keep-out zones, prohibited areas in Si for processing MOSFETS and other sensitive FEOL devices on IC layouts due to the underfill-microbump stress impact, are proposed. Circular keep-out zones, primarily above the microbump position are proposed for n-type devices. Rectangular keep-out zones around the microbump position, but not directly above the microbump position, are proposed for p-type devices.

Underfill-microbump stress mitigation guidelines are proposed and ranked according to benefiting impact, from highest to lowest impact on stress reduction in Si:

- increasing die thickness
- choosing a low-stress underfill
- decreasing the pitch between microbumps
- decreasing microbump height
- grouping microbumps in larger arrays

The underfill thermal shrikage defined by its CTE below and above its glass transition temperature has the highest impact on underfill thermo-mechanical behavior. Choosing an underfill with lower thermal shrinkage decreases stress levels in Si significantly. Additionaly, choosing an underfill with higher glass transition temperature and lower cure shrinkage could reduce stress levels in Si further. A process involving creation of etched rings in Si around the microbump is proposed in attempt to provide a potential

alternative to Si stress reduction. When grouping microbumps in an array and when microbump are at sufficiently close pitch for their stress fields to interact, an edge effect can be expected. The microbump array edge effect involves stress peaks in Si observed above microbumps on the edge of the microbump array. Otherwise, increasing the size of the microbump array decreases stress in Si above the microbumps within the microbump array.

FEOL stress after 3D IC stacking was further analyzed on FUJI and PTCQ stacks. The main differentiator between these new generations of stacks and the ETNA stack is an increase of top Si, from 25 μm to 50 μm. The PTCQ stacks additionaly decreased the microbump pitch from above 200 μm to 40 μm and grouped microbumps in large arrays over the whole stack.

In Si on the opposite side of the microbump tensile in-plane stresses are dominant. Directly above the projected microbump center, in-plane stress components are of equal value. P-type transistors in a [110]/[-110]/[001] reference frame will exhibit lower current shifts when placed directly above the projected microbump center as in that reference frame they have in-plane piezocoefficients of similar value and opposite sign. N-type transistors are preferable for monitoring the underfill-microbump stress mechanism in a [110]/[-110]/[001] reference frame on the opposite side of the microbump as their in-plane piezocoefficients are of the same sign. N-type transistors will therefore not cancel out the impact of in-plane stress components which are of the same sign and similar in that position. In Si below the microbump, n-type transistors could be more desirable as they exhibit a high sensitivity to out-of-plane stress. However, a combination of n-type and p-type sensors below the microbump enables individual in-plane and out-of-plane stress component extraction.

Underfill-microbump stress was analyzed on the FUJI stack on the opposite side of the microbump. Long channel p-FET arrays of the 32nm technology node with transistors oriented in longitudinal direction were placed in the vicinity of the microbump. A current shift of -6% was observed linked to 150 MPa of tensile stress. Current shift was also monitored at elevated tempertures up to 85 ºC. The impact of temperature on stress sensors was offsetted. Current shift decrease due to stress relaxation at higher temperatures was observed . Linear extrapolation of current shifts refered to an equivalent stress free temperature of the 3D IC stack just above 200 ºC. Finite element simulations show good agreement with observed current shifts at room temperature if the stack thermo-mechanical bonding temperature of 250 ºC is taken into account. According to electrical measurements on stress sensors and finite element simulations, stress build up above the glass transition temperature should not be neglected. More complex finite element models need to be built, based on more detailed studies of underfill behaviour above its T_g in order to extract the real equivalent zero stress temperature of a 3D stack.

The PTCQ test chip was designed to capture stress throughout the die after 3D IC stacking and packaging. N-type and p-type transistors oriented in longitudinal and transverse directions, individual and in array configuration, acted as global and local stress sensors. Current shifts were measured on the opposite side of the microbump, below the microbump and between the microbumps. The stress impact of the Cu pad above the FEOL was also analyzed. Stress extracted on the opposite side of the microbump equals to 125 MPa of tensile in-plane stress. Out-of-plane stress extracted

below the microbump equals to close to 400 MPa of compressive stress. In-plane and out-of-plane stress between the microbumps is below 100 MPa with out-of-plane stress lower than in-plane stress. On the thin Si die opposite of the microbump, where in-plane stresses are dominant, out-of-plane stress is negligible. The Cu pad induces non-negligible negative current shifts on n-type transistors below. Plastic deformation of Cu, most likely during thermo-compression bonding of the stack, could be the reason for remaining stress in Si below the Cu pad.

When stress in Si on the opposite side of the microbump is observed through the stack generations, a stress decrease is successfully observed, from an initial value of 827 MPa down to 150 MPa on the FUJI stacks and 125 MPa on the PTCQ stacks.

Chapter 8

2D and 3D packaging

Chapter 8 comprises results from studies on 2D and 3D IC packages. In total three package types were analyzed: a 2D BGA package with internal wirebonds, a 2D BGA with internal Cu pillars and a 3D SIC BGA package. All packages included the PTCQ test chip. The 2D IC packages were assembled for confirmation of mechanical stress principles in a 2D environment and as a test case for new methods of stress component extraction. The 2D packaging environment acted as a learning curve towards 3D SIC packaging. Section 8.1 compares the material properties of the used mold compounds and package substrate. Section 8.2 analyzes Si stresses in the 2D BGA package with wirebonds. Section 8.3 analyzes Si stresses in the 2D BGA package with Cu pillars, acting as a final step towards 3D packaging. Section 8.4 analyzes the stresses occuring after packaging of the PTCQ 3D stack.

Material properties	T_g [°C]	Cure Shrinkage [%]	E_1 ($T<T_g$) [GPa]	E_2 ($T>T_g$) [MPa]	CTE_1 ($T<T_g$) [ppm/°C]	CTE_2 ($T>T_g$) [ppm/°C]
EMC 1	134	0.31%	23	600	9	36
EMC 2	120	0.31%	22.5	15	10	40
EMC 3	130	0.22%	25	15	9	35
EMC 4	135	0.14%	26	15	7	29
EMC 5	150	0.04%	25.5	160	7	30
Substrate	200	-	28	-	14	-

Table 8.1: Material properties of 5 epoxy mold compounds (EMC) and the laminate substrate material that were used for 2D and 3D packaging

8.1 Impact of packaging in 2D technology

The PTCQ die was packaged in 2D IC and 3D SIC technology. The 2D IC packages were assembled for confirmation of mechanical stress principles in a 2D environment and as a learning curve towards 3D SIC packaging. Subsequently, 3D SIC packages were assembled and their thermo-mechanical impact analyzed. To this day, no published results on the thermo-mechanical behavior of 3D SIC packages was present.

In total 3 package configurations were assembled:

- 2D BGA package with internal wirebonds
- 2D BGA package with internal Cu pillars
- 3D SIC BGA package

Table 8.1 summarizes material properties of 5 epoxy mold compounds (EMC) and the laminate substrate material that were used for 2D and 3D packaging. The main differentiators of these mold compounds are their glass transition temperature and CTE above the glass transition temperature. The laminate has a high glass transition temperature, above the mold compound processing temperature.

8.2 2D wirebonded package

The first assembled package involved a 2D BGA package, illustrated in figure 8.1, where the PTCQ die FEOL, facing upwards, was attached to the package substrate with the die attach glue and the electrical connections were established by means of wirebonds. The thickness of the Si equaled to 300 μm, the substrate 240 μm and the total height of the mold compound 700 μm. Slightly less than 400 μm of mold compound was present above the die as the die thickness measures nominally 30 μm. The 2D BGA package with wirebonds was overmolded with two mold compounds, EMC 1 and EMC 2, properties of which are shown in table 8.1.

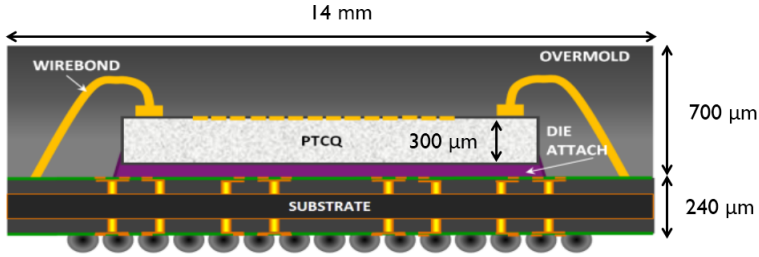


Figure 8.1: Illustration of the first assembled package with the PTCQ die. In the 2D BGA package the PTCQ die FEOL, facing upwards, was bonded to the substrate with wirebonds.

Figure 8.2 presents the first application of the package where in-plane stress extraction explained in section 5.2.3 was demonstrated. Global sensors, individual n-FETs in longitudinal and transverse direction from cell 1 were measured on wafer level and after packaging. When the FEOL is facing upwards towards the mold compound, out-of-plane stress can be considered negligible and in-plane stresses dominant. The stress map in Si on the active side facing the mold compound is similar to the biaxial stress test in section 5.4, figure 5.24. The shear and out-of-plane stresses become dominant at the edges of the test chip, below 100 μm distance from the die edge. In-plane stress can be extracted using one longitudinal and one transverse transistor of the same type according to eq. 5.7 and 5.8. The extracted in-plane stresses are additionaly considered equal across the die, far enough from the die edges. Figure 8.2 presents extracted in-plane stress values from several distributed points across the die indicated in the illustration above the results. Compressive stress is observed on the die surface, consistent with the shrinking nature of the mold compound after cooling from processing temperature of 180 ºC to room temperature. The stress values exhibit a considerable variation with most stress values around -100 MPa. Higher values are assumed to occur due to local filler particles of the mold compound in the vicinity of the transistor pressing down on the trasistor location during cooling. The instability of the results might be connected to the very bad yield of the packages. In several cases, the whole package or parts of the package were not electrically functional.

Further investigation was done on the warpage of the 14mm x 14mm package with EMC 2 and comprising 8mm x 8 mm die with optical profilometry, mechanical profilometry and X-ray, presented in figure 8.3. Mechanical and optical profilometry was performed on the package mold compound surface and revealed a convex shape over the package with small concave regions towards the package edges, figure 8.3 a). Intuitively, if the convex shape is transfered to the die as well, this would indicate tensile stress on the die surface which is in conflict with the electrical measurments. X-ray diffraction was employed to investigate the warpage of the comprising die. These measurements were performed by the team of prof. P.J. McNally of DCU. X-rays penetrate to the desired top Si surface and reveal warpage due to the shifts of the diffracted Bragg angle. The trend of the diffracted Bragg angle can indicate convex or concave warpage. A rocking curve (RC) map of the top Si surface was plotted in figure 8.3 b) indicating

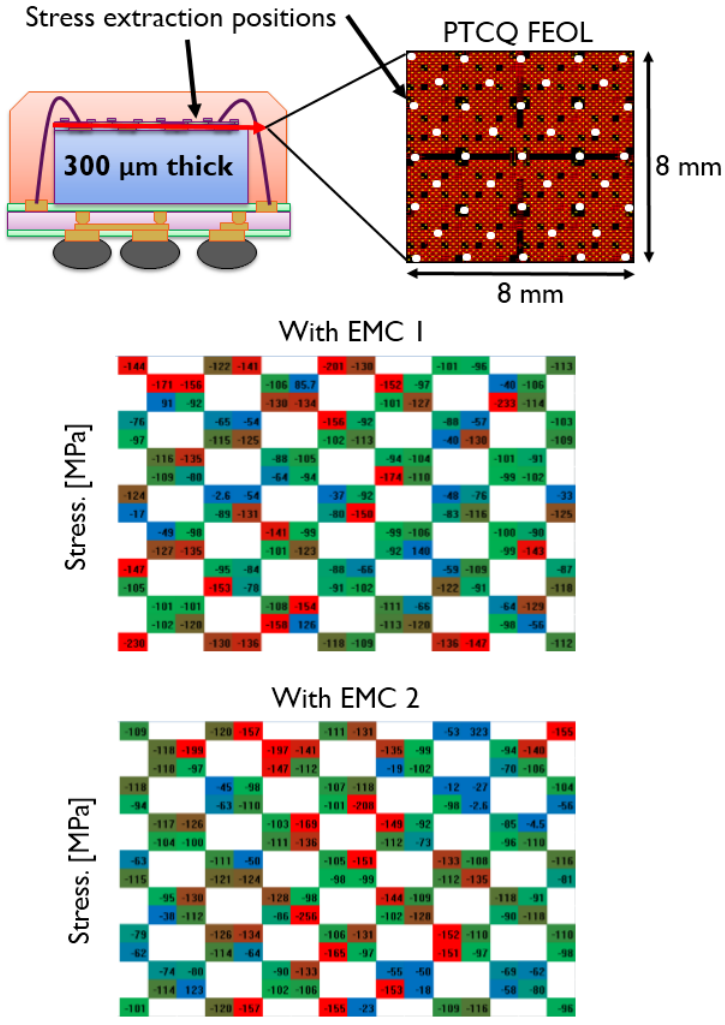


Figure 8.2: Demonstration of in-plane stress extraction from the PTCQ die after 2D packaging, in case of EMC 1 and EMC 2

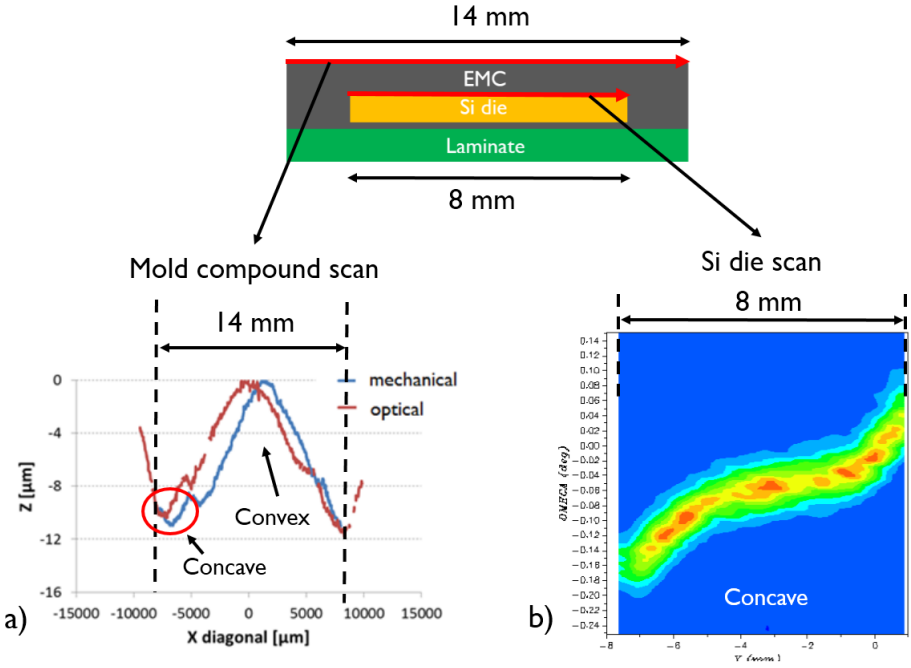


Figure 8.3: Further investigation was done on the warpage of the 14mm x 14mm package with EMC 2 and the warpage of the 8mm x 8 mm die inside the package with: a) optical profilometry and mechanical profilometry and b) X-ray diffraction. Profilometry indicates a convex shape of the package while X-ray reveals the inner die warped in a concave shape.

concave nature of Si die top surface warpage. Therefore, the mold compound surface above the top Si die surface exhibits convex warpage with small concave regions on the edges of the package while the Si die within the package exhibits concave warpage.

The following hypothesis to explain the occuring warpages is proposed. Bonding of the die to the laminate substrate with the die attach is usually done at temperatures equal to or slightly below the mold compound processing temperature of 180 ºC. The die is placed on the viscous die attach at room temperature. While raising the temperatures, expansion of both the Si and substrate occurs, however bending of the structure does not occur as the die attach is in liquid state, unable to create a bending moment to the die to substrate system. Therefore, at 180 ºC or slightly lower, the die to substrate structure can be assumed as flat. Following the hardening of the die attach after prolonged exposure to high temperature, a firm die to substrate link is created and during cooling to room temperature, bending of the structure can be expected, most likely in convex shape as the CTE of the laminate is higher than the one of Si. The mold compound is dispensed again at 180 ºC. At this temperature, the die to substrate structure can again be assumed to be flat. In case the mold processing temperature is higher than the previous die to substrate bonding temperature, a slight convex die to substrate

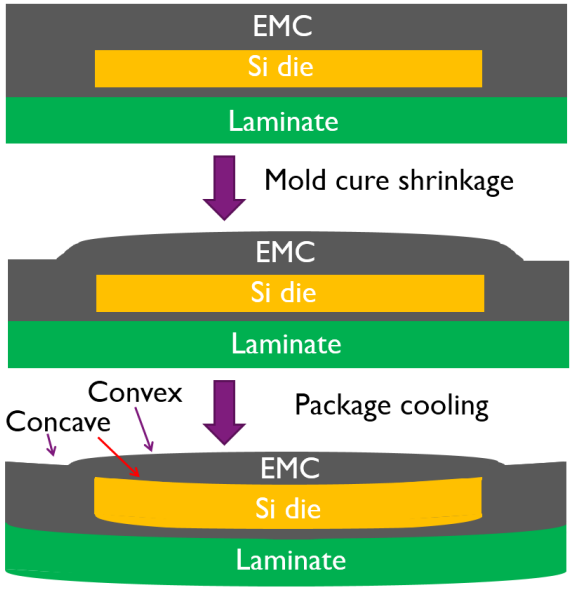


Figure 8.4: Assumed mechanism resulting in warpages observed in figure 8.3. At the mold process temperature of 180 ºC, during mold compound hardening, the shrinkage of the mold compound caused by material curing results in a bubble of mold compound material above the die. In essence, the mold compound shrinks during curing and starts feeling and impact from the shape of the Si die finally creating convex curvature above the Si die. During cooling of the package, due to higher thermal shrinkage than the substrate, the mold compound acts in a concave way, which could explain the concave observed shape at the edges of the package in figure 8.3 a).

shape might occur. After mold hardening at 180 ºC, during cooling of the package, the shrinking of the mold compound is dominant. The calculated thermal shrinkage of the mold compound, given by eq. 7.2, equals to 0.33% while the substrate's thermal shrinkage equals to 0.22%. From the point of thermal shrinkages of materials within the package, the package should have an overall concave shape. An additional effect is assumed to play a role, presented in figure 8.4. At the mold process temperature of 180 ºC, during mold compound hardening, the shrinkage of the mold compound caused by material curing results in a bubble of mold compound material. In essence, the mold compound shrinks during curing and folds around the Si die creating convex curvature above the Si die. During cooling of the package, the mold compound initiates a concave warpage of the package, which could explain the concave observed shape at the edges of the package in figure 8.3 a).

8.3 2D BGA

The following 2D package included a BGA with Cu pillars between the die and laminate substrate, as illustrated in figure 8.5. The thickness of the die in this case was 200 μm, the substrate 390 μm, the Cu pillars spanned 50 μm in diameter and 70 μm in height while the mold compound extended to 770 μm. The package planar dimensions remained at 14mm x 14mm.

The package was assembled with 3 different mold compounds, EMC 1, EMC 2, EMC 3 and without a mold compound, bonding only to the substrate. Global sensors in cell 1, both n-FET and p-FET longitudinal and transverse devices were measured on 9 different positions on the diagonal of the die. The illustrated measurement positions across the die and the observed current shifts from the FETs are summarized in figure 8.6. Both n-FETs and p-FETs reveal the same current shifts in longitudinal and transverse directions. As was the case with global sensors and microbumps after 3D stacking in section 7.4.5.1, the global sensors here after 2D packaging are placed in equal distance from the surrounding Cu pillars. The equal current shifts in the two different orientations confirm that the two in-plane stresses can be considered of the same value. It is important to notice that the case without mold compound exhibits similar ranges of current shift as the cases with mold compounds. P-type transistors exhibit a greater difference between all cases with the die to substrate case exhibiting the highest current shift.

Assuming equal in-plane stress components and utilizing one n-type transistor in transverse orientation and one p-type transistor in transverse orientation, according to eq. 7.10 and 7.11, in-plane and out-of-plane stress can be extracted. Figure 8.7 summarizes the stress extraction results on the same 9 positions across the die, with 3 different mold compounds and in absence of a mold compound, with only the substrate attached. The structure with only the substrate attached to the die exhibits the highest in-plane stress with slightly higher than 200 MPa of compressive stress. Adding mold compounds lowers the stress. For EMCs 2 and 3, between 150 MPa and 200 MPa of compressive stress and for EMC 1 around 100 MPa of compressive stress. The out-of-plane stress varies for all cases between 50 MPa of compressive and 50 MPa of tensile stress in Si. When the substrate is attached, compressive stress is visible, while adding

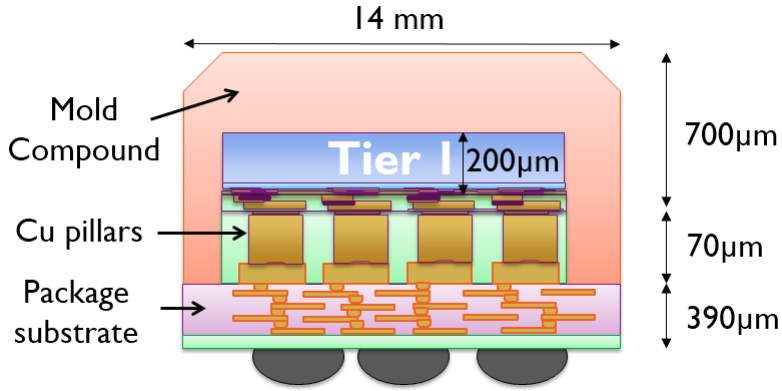


Figure 8.5: Illustration of the second assembled package with the PTCQ die. In the 2D BGA package the PTCQ die FEOL, facing downwards, was bonded to the substrate with Cu pillars.

mold compounds changes the sign to tensile stress.

Optical profilometry was performed on the die in case of attachment to substrate, shown in figure 8.8. The curvature on the top die surface indicates convex shape of the die. This shape of the die implies tensile stress on the back side of the die and compressive stress on the FEOL side. After bonding to the substrate at elevated temperatures, the higher CTE of the substrate causes convex bending of the die to substrate structure at room temperature. The die to substrate structure is then heated again to 180 ºC correcting somewhat the curvature. However, in the BGA package with Cu pillars, attachment to substrate is performed with thermo-compression bonding done at temperatures of 250 ºC, higher than mold compound processing. This means that after heating the die to substrate structure to mold processing temperature of 180 ºC the structure is likely still in convex shape. After mold processing, during cooling of the package, the mold compound with a higher thermal shrinkage than the substrate becomes dominant and attempts to correct the curvature, however certain levels of compressive stress remain observed as a negative current shift of n-FETs and positive current shift of p-FETs. More convex curvature is created than the mold compound can correct as die to substrate assembly was done at temperatures 70 ºC higher than the mold compound processing.

8.4 Stress build up in a 3D stacked IC package

Following 2D packages, first 3D packages were processed. These included packaging the PTCQ stack from section 7.4.5. The PTCQ 3D BGA package is illustrated in figure 8.9. Instead of Cu pillars, flip chip solder balls of 70 μm in height were used. Two substrate thicknesses were used, 280 and 390 μm. The mold compound thickness of 700 μm and 14mm x 14mm package dimensions remained the same as in the 2D BGA case. Figure 8.10 shows photographs of the 3D SIC package, where figure 8.10

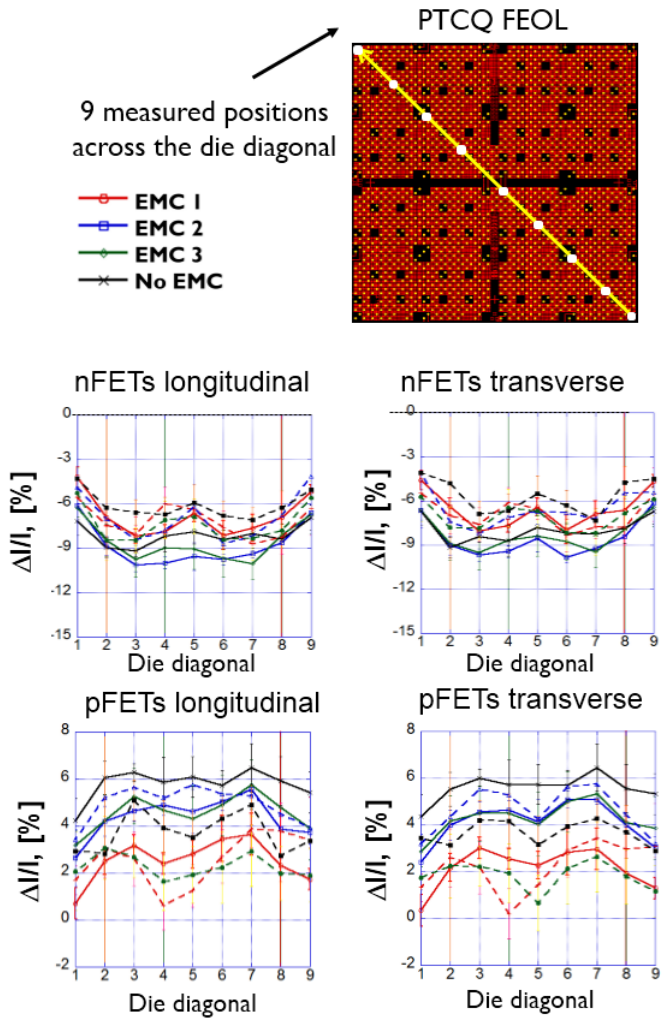


Figure 8.6: Current shifts observed on the n-FET and p-FET global sensors taken from 9 measurement positions across the die diagonal

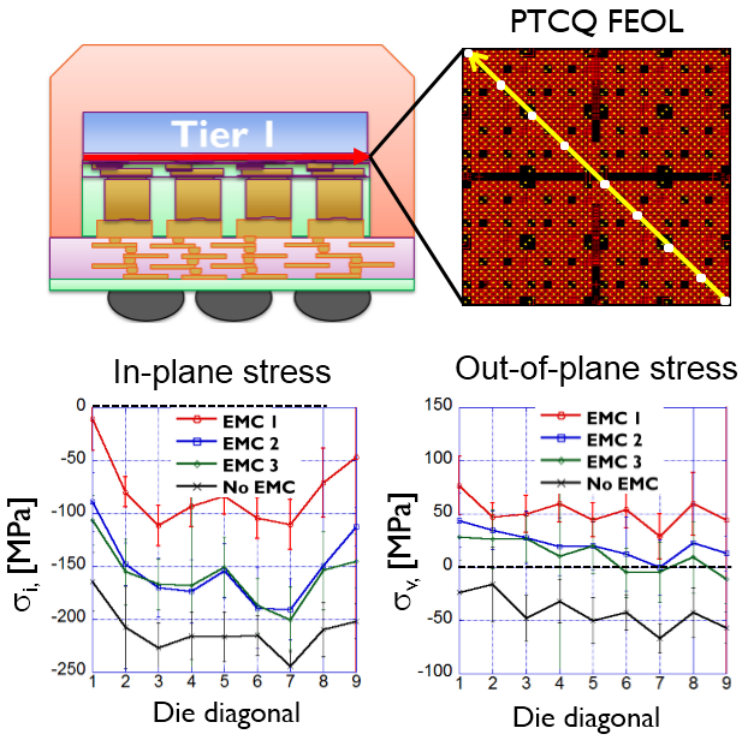


Figure 8.7: Extracted in-plane and out-of-plane stress on the same 9 positions across the die diagonal, with 3 different mold compounds and in absence of a mold compound, with only the substrate attached. The structure with only the substrate attached to the die exhibits the highest in-plane stress.

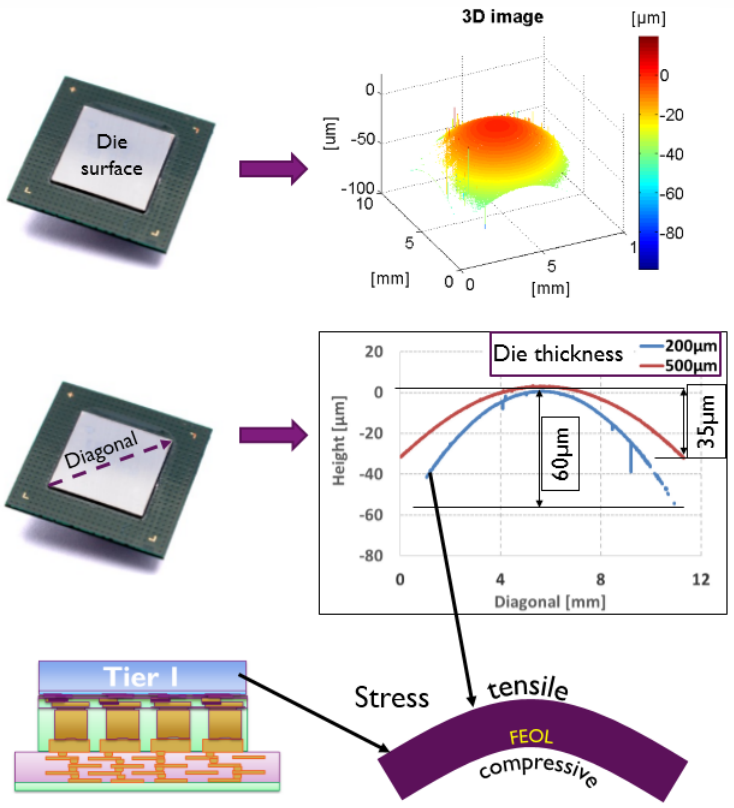


Figure 8.8: Optical profilometry performed on the die in case of attachment to substrate. The curvature on the top die surface indicates a convex shape of the die. This shape of the die implies tensile stress on the back side of the die and compressive stress on the FEOL side.

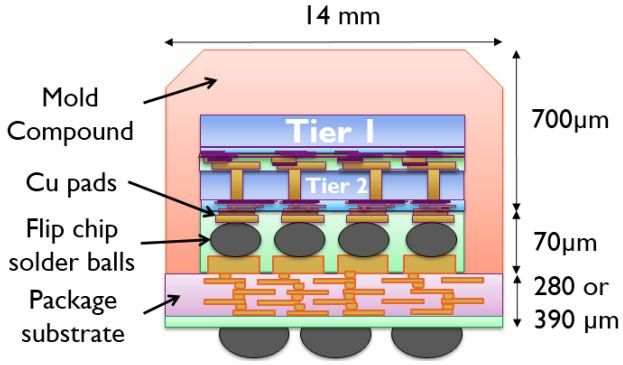


Figure 8.9: Illustration of the assembled 3D SIC package with the PTCQ stacks. The stack is flipped with the PTCQ FEOLs facing downwards. Tier 2 is connected to the substrate with flip chip solder balls instead of Cu pillars.

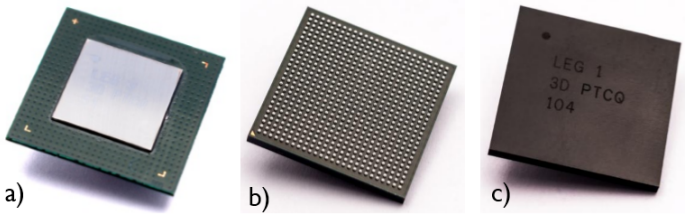


Figure 8.10: Photographs of the 3D SIC package, where figure 8.10 a) reveals the 3D stack to substrate assembly phase, figure 8.10 b) the backside of the package with the solder ball array visible and figure 8.10 c) the top side of the molded package.

a) reveals the 3D stack to substrate assembly phase, figure 8.10 b) the backside of the package with the solder ball array visible and figure 8.10 c) the top side of the package. Global and local sensors were measured across the die to capture the impact of packaging the PTCQ 3D stack. The reference value for current shifts were the 3D stack values. In total 4 configurations were processed. Three of them, packages with EMC 2, with EMC 4 and a bare 3D stack were assembled on a thicker substrate of 390 μm. The thinner substrate of 280 μm was used to assemble the packages with EMC 5.

Figure 8.11 summarizes the obtained current shifts from global sensors, n-type and p-type longitudinal FETs from the 9 positions across the die, on both Si tiers. N-type transitors on both tiers exhibit a negative current shift while the p-FET transistors exhibit a negative current shift on tier 1 and positive current shift on tier 2. The die to substrate structure continues to exhibit stresses within range of the full packages with mold compounds. While packages with EMC 2 and EMC 4 seem to exhibit similar levels of stress, the package with EMC 5 and a thinner substrate seems to separate itself from the other configurations induce highest current shift on n-type FETs, highest current shift on tier 2 p-FETs and close to zero current shift on tier 1 p-FETs.

If the in-plane components of stress in positions of global stress sensors are taken as equal, due to their position between microbumps and Cu pads, using eq. 7.10 and 7.11, in-plane and out-of-plane stress can be extracted. In this case, instead of two transverse oriented FETs of the same type, longitudinal FETs of opposite type are used. Figure 8.12 summarizes the extracted in-plane stress and out-of-plane stress values in position of global sensors for tier 1 and tier 2. Due to the fact that the FET currents were referenced to 3D stack level FET currents and not wafer level FET currents, the impact visible is solely from the impact of packaging - substrate assembly and mold compound processing. According to results from figure 8.12, 3D packaging induces compressive in-plane stresses on tier 2 closer to the substrate and tensile in-plane stresses on tier 1 surrounded by the mold compound. Furthermore, out-of-plane stresses on both tiers appear to be tensile with significantly higher out-of-plane stress on tier 1.

Considering the impact of the 3D stack attached to the package substrate without processed mold compound, optical profilometry was performed on the top surface of tier 1, opposite of the tier 1 FEOL position, presented in figure 8.13. A convex shape of the 3D stack was observed which would indicate compressive stresses in the FEOL region of the top tier. The substrate material shrinks the most during cooling from the reflow bonding temperature of 250 ºC due to its high CTE value and bends the above 3D stack. Application of the mold compound, similar to the 2D BGA case, attempts to change the warpage of the stack, therefore also impacting stress changes. Tier 2 closer to the substrate, figure 8.12, exhibits similar stress trends as the die in a 2D BGA package, figure 8.7. After packaging compressive in-plane stresses and tensile out-of-plane stresses are observed on tier 2. On tier 1, which back side is in contact with the mold compound, in-plane stress changes sign and out-of-plane stress increases significanlty.

It can be asssumed that this change can be correlated to a more dominant impact from the mold compound than the substrate. One hypothesis might be that during shrinking of the whole package, the shrinking mold compound tries to counteract the effects of the substrate, having more success on tier 1 than tier 2. The shrinking mold compound prevails on tier 1 causing concave warpage of tier 1 leading to tensile in-

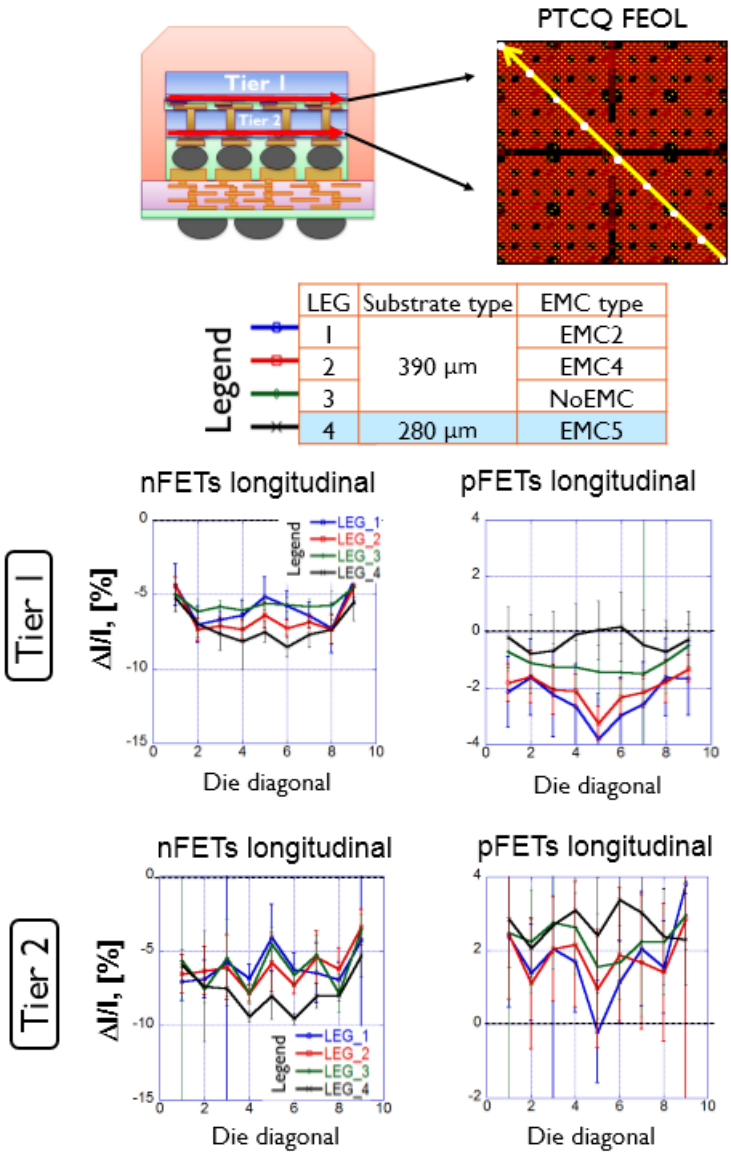


Figure 8.11: Current shifts obtained from global sensors, n-type and p-type longitudinal FETs from the 9 positions across the die diagonal, on both Si tiers

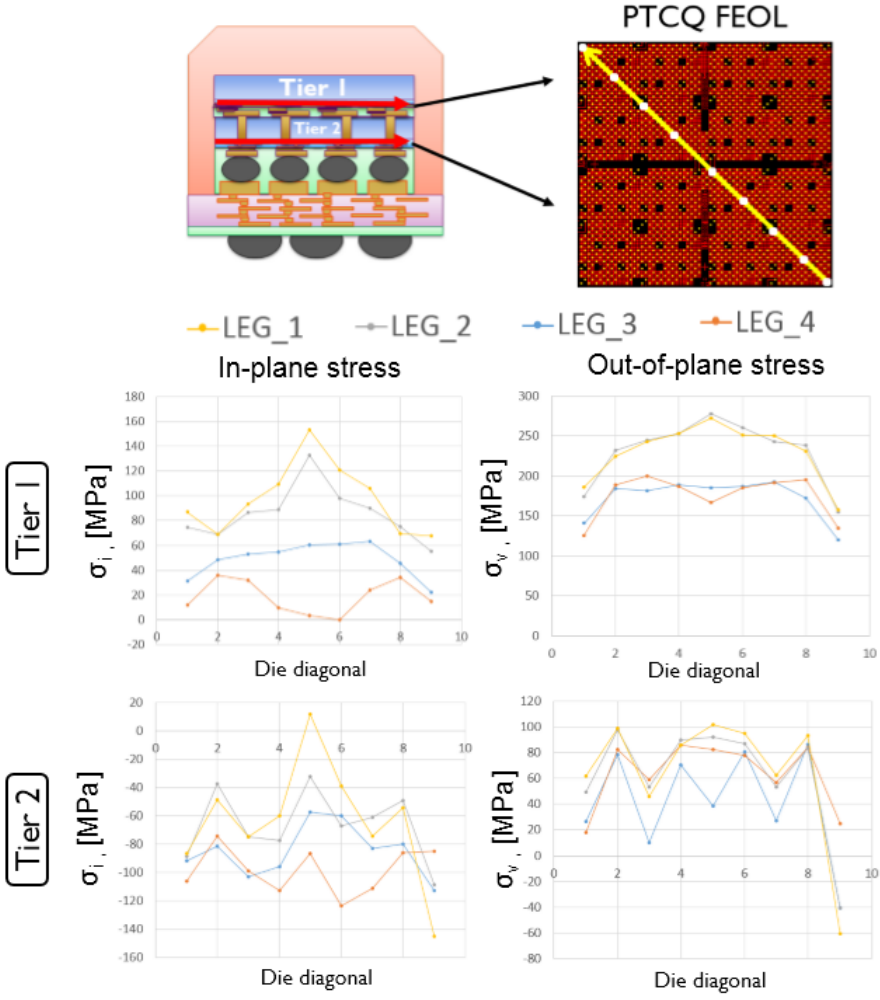


Figure 8.12: Extracted in-plane stress and out-of-plane stress from the global sensors for tier 1 and tier 2

plane stresses in the FEOL. The stress trends between different configurations remain as on tier 2 but are shifted towards the tensile region between 100 and 150 MPa. The shift in out-of-plane stress is more ambigous. High out-of-plane stress might be created due to local shrinking of the underfill, acting in the opposite direction of the mold compound. While the mold compound pulls the die upwards, the underfill pulls it downwards. In case of the 2D BGA package, figure 8.7, addition of the mold compound cause a shift from compressive to tensile out-of-plane stress, perhaps due to the mentioned die pulling. Tensile out-of-plane stress is visible on both tiers in figure 8.12 as well, however the out-of-plane stress on tier 1 closer to the mold compound is significantly higher. Out-of-plane stress values on tier 2, figure 8.12 are comparable to the molded package out-of-plane stress values in figure 8.7. A solid hypothesis as to why the out-of-plane stress on tier 1 is considerably higher still needs to be set. One point worth investigating is the thermo-mechanical properties of the underfill between the dies and the underfill between tier 2 and the substrate. The stack underfill might be pulling stronger than the die to substrate underfill creating higher out-of-plane stress between the dies.

Figure 8.14 summarizes the electrical measurements after packaging on the local n-FET longitudinal sensors. The values before packaging are current shifts obtained on 3D stack level. These current shifts are originally referenced to wafer level measurements, as presented in figure 7.53. In figure 8.14, the current shifts are referenced to the minimum value within the n-FET array for better visibility of the current shifts occuring after 3D packaging. The microbump n-FET arrays on tier 1 and tier 2 for EMC 2, EMC 4 and no EMC exhibit negative current shifts of approximately -8%. This is in range with the observations of the global n-FET sensors as well. The configuration with a thinner substrate and EMC5, as with the global sensors, exhibits slightly higher current shifts of -10%. The shapes of microbump and Cu pad impact are preserved after 3D packaging.

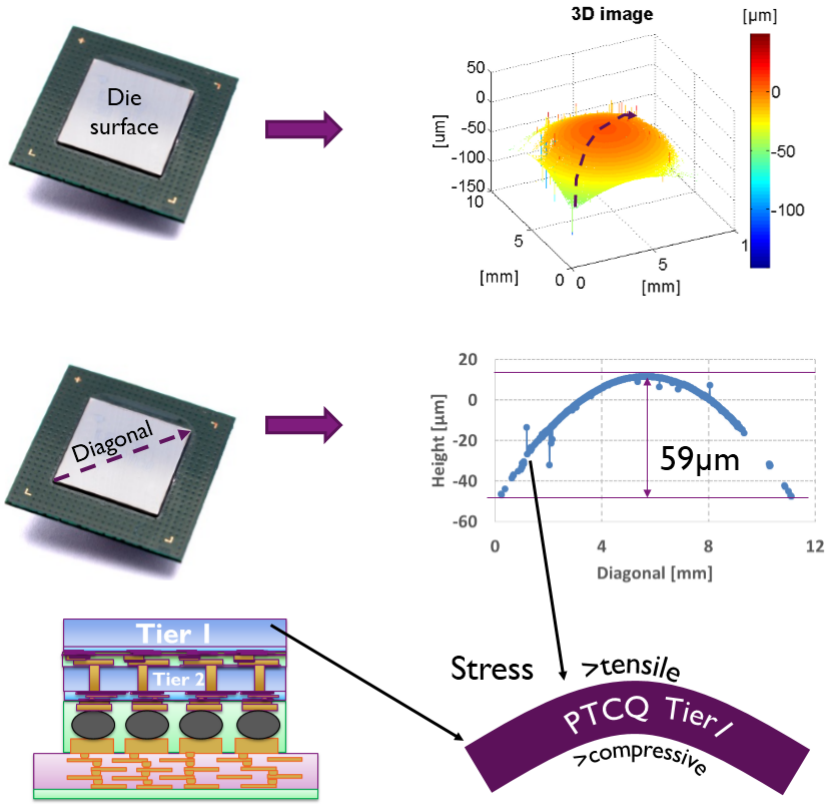


Figure 8.13: Optical profilometry was performed on the top surface of tier 1, opposite of the tier 1 FEOL position indicating convex warpage

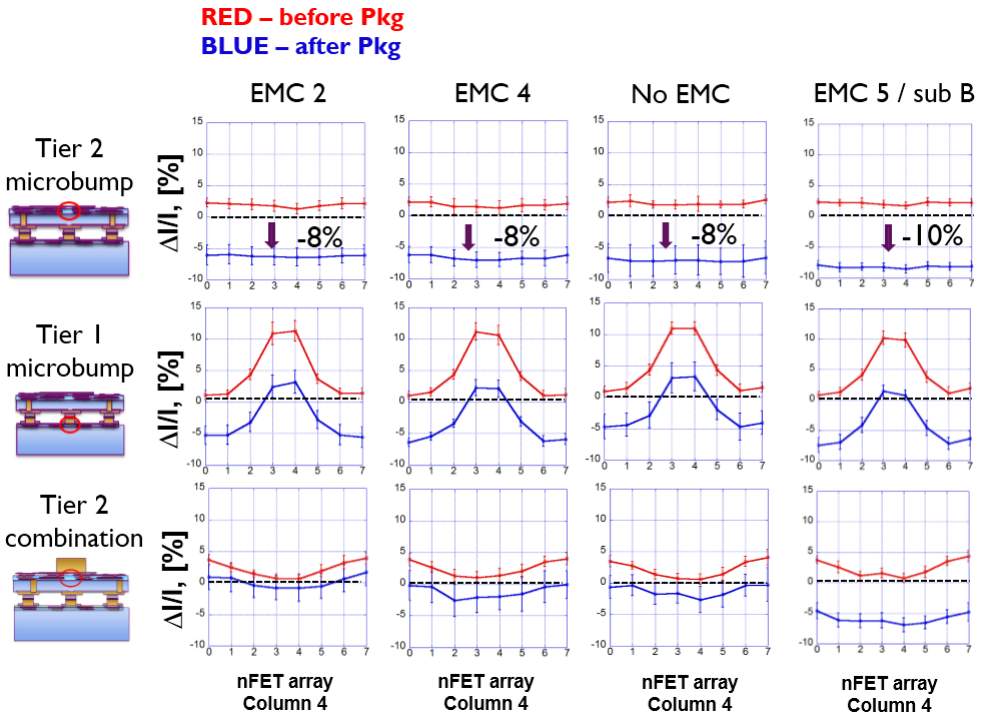


Figure 8.14: Current shifts observed after packaging on the local n-FET longitudinal sensors. The values before packaging are current shifts obtained on 3D stack level. These current shifts are originally referenced to wafer level measurements, as presented in figure 7.53. The current shifts are referenced to the minimum value within the n-FET array for better visibility of the current shifts occuring after 3D packaging.

8.5 Summary

Stress is analyzed in 2D and 3D SIC packages. In total 3 package types were assembled with the PTCQ test chip:

- 2D BGA package with internal wirebonds
- 2D BGA package with internal Cu pillars
- 3D BGA package

The packages were processed with 5 different mold compounds. Also assemblies of a single die on substrate and 3D stack on substrate, without mold compound, were made. The main difference between the substrate and mold compounds is the higher CTE of the mold compound above its glass transition temperature.

In-plane stresses in Si were extracted from MOSFET current shifts in the 2D BGA package with internal wirebonds. The FEOL of the Si die was in this case facing the mold compound. Compressive stress around 120 MPa was observed. High stress is observed in random locations on the die possibly linked to mold compound filler particles pressing on the die during mold compound cooling. Convex warpage of the package surface was observed with optical and mechanical profilometers. X-ray revealed concave warpage of the inner Si die. The convex shape of the package is linked to the mold compound cure shrinkage during mold compound hardening. The concave shape of the die is linked to the CTE mismatch between the mold compound and the substrate. After the convex shape of the mold compound is created due to mold compound cure shrinkage at the mold compound processing temperature of 180 ºC, during cooling of the package to room temperature, due to higher thermal shrinkage of the mold compound, concave warpage is emposed on the stack. Convex warpage remains on the outer side of the package above the die, while concave warpage impact is observed on the package edges further from the die where only mold compound and substrate interact.

In-plane and out-of-plane stresses in Si were extracted from MOSFET current shifts in the 2D BGA package with Cu pillars. The FEOL of the Si die was in this case facing the Cu pillars and substrate. Compressive in-plane stress and out-of-plane stress were observed after die bonding to the substrate. Compressive, but lower in-plane stress and slightly tensile out-of-plane stress is observed after additional processing of the mold compound. Optical and mechanical profilometry were performed on the die to substrate structure exhibiting convex die warpage. It is assumed that during cooling of the die to substrate structure, the substrate with higher CTE than Si shrinks and causes a convex warpage of the structure. Since the die FEOL is on the substrate side, convex warpage on the back Si die surface points to tensile stress on the back Si die surface and compressive stress on the inner Si die surface, where the die FEOL is. This is consistent with the extracted stress values from electrical measurements. When the mold compound is applied, the higher thermal shrinkage of the mold compound compared to the substrate, as in the case of the 2D BGA package with internal wirebonds, attempts to create concave warpage, opposite of the effect of substrate assembly. However, in the case of the 2D BGA with Cu pillars, the die is attached with the substrate

by thermo-compression bonding at 250 ºC, higher than the mold compound processing temperature. This means that at mold compound processing temperature of 180 ºC, the die to substrate structure is probably already bent in a convex shape and the mold compound does not succeed in full correction or change of the warpage. This is why the in-plane stresses are still compressive after adding the mold compound but lower than in the die to substrate structure case. The substrate can therefore have a stronger impact on stress in Si than the mold compound. Furthermore, the mold compound could be used to correct the package warpage and to an extent lower Si stress.

In-plane and out-of-plane stresses in Si were extracted from the 3D IC package from MOSFET current shifts from both Si tiers, from local and global sensors. Flip chip solder balls were used instead of Cu pillars. Regarding global sensors, placed between the microbumps, on tier 2, the bottom die in the package with its FEOL facing the substrate, compressive in-plane stress around 100 MPa and tensile out-of-plane stress below 100 MPa was extracted. Optical and mechanical profilometry confirm a convex warpage of the 3D IC stack when attached to the substrate. This situation is comparable to the 2D BGA package with Cu pillars, where the FEOL is also facing downwards towards the substrate.

Different stress trends are observed on tier 1, the die further from the substrate, on the mold compound side. This die FEOL is facing the underfill-microbump region and the die it is stacked on while the back side of the die is turned to the mold compound. There, tensile in-plane stress and higher tensile out-of-plane stress than on tier 2 occurs. It is known from 2D BGA packages, that the mold compound tries to bend the package in a manner opposite of the substrate due to the higher thermal shrinkage of the mold compound compared to the substrate. It can be argued that the mold compound has a higher and more direct impact on the die further away from the substrate and therefore manages to create concave warpage of the top Si die resulting in such stress fields. Although the substrate undoubtably plays a role in the overall warpage, the mold compound does not interact only with the substrate but with the bottom Si die of the 3D IC stack as well. In one approximation, the bottom Si die can be viewed as the substrate. In terms of warpage of tier 1, in that case, two interactions are occuring. The mold compound to substrate interaction and the mold compound to the bottom die interaction.

Local sensors were also analyzed in the 3D IC package. Current shift shapes observed from the underfill-microbump mechanism are preserved and globally shifted by -8% after packaging. The stress shapes and local current shifts on tier 1 are preserved regardless of the mold compound impact observed on the global sensors as the out-of-plane stress caused by the underfill-microbump mechanism plays a dominant role.

Chapter 9

General conclusions and future work

9.1 General conclusions

9.1.1 Initial objectives and final contributions

The objective of this PhD thesis was to detect, describe and minimize the negative thermo-mechanical effects of 3D IC stacking and packaging on the FEOL in order to improve 3D IC stack/package performance. This objective was channeled into the following main goals:

- Develop a sensor-based CPI methodology for the FEOL
- Provide understanding and mitigation guidelines on the generated mechanical stress to the FEOL in the vicinity of the microbumps after 3D stacking - refered to as the underfill-microbump stress mechanism

Additional goals developed throughout the thesis included:

- Enable benchmarking of FEOL device stress sensitivities from scaling technology nodes by applying stress calibration methodology used for stress sensors
- Monitor and interprete stress evolution after 3D IC stacking and 3D IC packaging

The forementioned goals have been met. With respect to the presented overview of this work in previous chapters, a list is given with the contributions of this thesis to the scientific community:

- First electrical measurements revealing the underfill-microbump impact on FEOL in 3D IC stacks
- Provided understanding on the underfill-microbump impact to the FEOL through experimental work and finite element simulations
- First proposed underfill-microbump stress mitigation guidelines concerning its impact on the FEOL
- Proposed an evaluation methodology for stress sensors for monitoring CPI in 3D IC stacks
- Demonstrated in-plane and out-of-plane stress calibration methods for FEOL devices
- Evaluated, selected and implemented MOSFETs as stress sensors that monitored CPI effects in 3D stacks and 3D packages: the underfill-microbump stress mechanism, impact of Cu pads and the interaction between the package substrate and mold compound
- Analyzed stress in Si in 3 different positions in the stack: on the opposite side of the microbump, below the microbump and between microbumps
- Extracted and interpreted individual in-plane stress and out-of-plane stress components in Si in a two layer 3D IC stack and 3D IC package

A large part of this work is a result of teamwork with particular contribution from Dr. Vladimir Cherman regarding test setup and measurements. The general conclusions of the thesis follow.

9.1.2 Evaluation of FEOL devices as CPI stress sensors

Regarding stress sensors in Si dies, this thesis focused on exploitation of MOSFETs. Applicability of FinFETs and pseudo-Hall transistors was also assessed. The linear piezoresistance model, originally developed for bulk Si, was applied to describe the stress behavior of FEOL devices. In this case, the piezocoefficients represent the stress sensitivity of the whole device rather than just bulk Si. For Si based FEOL devices, such as MOSFETs, their drain to source current is directly proportional to electron or hole mobility. The piezoresistance model applied to FEOL devices linked independent mechanical stress components, over device piezocoefficients, with device current shift.

A stress sensor evaluation methodology was proposed and implemented. Its 3 phases are wafer level evaluation, stress calibration and sensor usage. During wafer level evaluation, the stability of the device is assessed with no external loads applied. If the device exhibits sufficient electrical stability, it can proceed to the stress calibration phase. Stress calibration involves current shift monitoring during application of a controlled external mechanical load. Device piezocoefficients can be extracted as a result of stress calibration which describe the sensitivity of the device to in-plane, out-of-plane or shear stress. If the device exhibits sufficient sensitivity to stress it can be submitted to a validation test. The validation test, where possible, is the final step before sensor application in a real environment. The Si die with processed stress sensors is placed in a simplified and known stress environment. Stress is then extracted with stress sensors and compared to other stress extraction methods using techniques such as profilometry or finite element modeling.

The usage of tools in this work is summarized.

Usage of finite element analysis:

- predict mechanical stress in Si after 3D IC stacking and 2D and 3D packaging
- by utilizing test obtained transistor piezocoefficients, transfer simulated stress to transistor current shift
- compare with other stress extraction techniques
- upon finite element model validation with other stress extraction techniques, provide guidelines for stress reduction in Si

Usage of the delaminator:

- as a 4-pt bending tool to apply controlled in-plane stress during stress sensor calibration
- to apply controlled out-of-plane stress during stress sensor calibration

Usage of the nano-indenter:

- as an alternative to the delaminator for applying controlled out-of-plane stress during stress sensor calibration

Usage of profilometers:

- to determine the warpage of a Si die, 3D IC stack or package
- by obtaining the curvature of a die, using the Stoney equation to calculate stress in Si

Usage of X-ray diffraction:

- to determine warpage of packaged Si dies

Stress sensitivity of several FEOL device types and through several FEOL technology nodes was assessed. The stress sensitivities are analyzed from two perspectives:

- To gain information on the nature of stress sensitivity of FEOL devices and its trends through scaling technology nodes
- Their applicability for usage as FEOL CPI stress sensors

Stress sensitivities of MOSFETs, FinFETs and pseudo-Hall transistors were analyzed. Calibrations were performed in [110]/[-110]/[001] and [100]/[010]/[001] Si reference frames. 4-pt bending was used for application of in-plane stress. Out-of-plane stress was applied with a delaminator tip over a Si cube to the Si die and alternatively with the nano-indenter.

In-plane stress piezocoefficients are obtained from the following devices:

- n-type and p-type MOSFETs, short and long channel, from 3 technology nodes - 130nm, 65nm and 32 nm
- n-type and p-type FinFETs of one channel size from the 32nm technology node
- n-type and p-type pseudo-Hall transistors of one channel size from the 65nm technology node

Out-of-plane stress piezocoefficients are obtained from the following devices:

- n-type MOSFETs, long channel, from the 65nm technology node

MOSFETs, FinFETs and pseudo-Hall devices exhibit a sensitivity to mechanical stress. Smith piezocoefficients for bulk Si are not meant to be used for FEOL devices. Obtained FEOL device piezocoefficients can greatly differ from Smith piezocoefficients for bulk Si. Direct calibration of FEOL devices to stress is necessary.

Summary of in-plane stress calibrations:

- MOSFETs
  - n-type and p-type MOSFETs through all tested technology nodes exhibit sensitivity to stress

  - taking into account their stress sensitivity levels, they are good candidates for usage as CPI stress sensors
  - in the used reference frame, [110]/[-110]/[001], n-type in-plane piezocoefficients are of the same sign and different magnitude, p-type piezocoefficients are of opposite sign and similar magnitude
  - a decrease of sensitivity with newer technology nodes is observed
  - long channel transistors exhibit higher sensitivity than short channel transistors
  - A MOSFET operating in the saturation regime is less senstive to stress than when operating in the linear regime
  - Stress sensitivity of MOSFETs is highly dependent on temperature and drops with temperature increase
  - MOSFETs in a biaxial stress environment reacted equivalently to a superposition of uniaxial stress environments
  - The linear piezocoefficient model might be invalid for high values of stress, approximately above 1 GPa due to non-linear relation between stress and current shift

- FinFETs
  - FinFETs exhibit similar sensitivity in longitudinal direction, but have a distinctively low stress sensitivity in the transverse direction
  - Due to the lower stress sensitivity in transverse direction, they are not recommended for CPI sensors
  - Lower transverse stress sensitivity is most likely not attributed to lower stress transfer in the transverse direction but a fundamental response of FinFET side walls

- Pseudo-Hall transistors
  - Pseudo-Hall transistors present an interesting alternative to transistor current shift monitoring for stress observation, by measuring voltage perpendicular to the drain and source transistor contacts
  - their stress sensitivity is independent of temperature, however the absolute value of the voltage measured to monitor stress still changes with temperature
  - Pseudo-Hall transistors did not exhibit an additional advantage over MOSFETs for implementation as CPI Si stress sensors in 3D IC stacks and packages due to:
    * The stress sensitivity invariance with temperature is not a particular advantage as stress from 3D IC stacks and packages is extracted at room temperature. They could provide more or arguably better performance than MOSFETs for stress extraction at higher temperatures

    * Extraction of individual stress components with pseudo-Hall transistors is not straightforward

Summary of out-of-plane stress calibrations:

- MOSFETs
  - n-type transistors exhibited higher sensitivity to out-of-plane stress in [001] direction than to in-plane stress in [110] and [-110] direction
  - p-type transistors were not calibrated to out-of-plane stress, according to Smith's bulk Si piezocoefficients, p-type Si on [001] Si surface exhibits low sensitivity to out-of-plane stress

An important advantage of MOSFETs is its ability to extract individual stress components. In certain conditions, MOSFETs can be used to extract both in-plane stresses and out-of-plane stress. Shear stress calibration was not attempted on any of the devices. According to piezoresistive theory, Si is not affected by shear stress in [110] and [-110] direction. Mechanical dicing reduces the ultimate tensile strength of Si. Mechanical polishing of Si sample edges did not result in an increase of its ultimate tensile strength. Ion etching of Si samples provided promising results with sample reaching GPa range without breaking. This could enable FEOL device stress calibration at GPa stress levels.

9.1.3 Analysis of mechanical stress in 3D stacked ICs

High current shifts above 40% were initially observed on n-FET transistors within a 3D IC stack, placed on the thin Si die, on the opposite side of the microbump. The initial stack consisted of a 25 μm thick die stacked on a 550 μm thick die. This current shift was linked to stress generated by the underfill-microbump stress mechanism. During cooling of a 3D IC stack after die bonding, the underfill as the material with the highest CTE in the surrounding, shrinks and pulls the thin Si die over the underlying microbumps causing local warpage of the thin Si die. The underfill microbump stress mechanism is a CTE mismatch driven mechanism initiated by the underfill material. The 40% current shift was linked to 827 MPa of Si stress.

Generated stress patterns in Si from the underfill-microbump mechanism on the opposite side of the microbump are circular. Negligible amounts of stress are observed in Si if an underfill is not present.

N-type and p-type Si and in that sense n-type and p-type MOSFETs have a distinctively different response to mechanical stress. The circular stress in Si on the opposite side of the microbump generated by the underfill-microbump mechanism creates circular, dominantly positive current shift patterns for n-type Si and orbital current shift patterns for p-type Si with regions of negative and positive current shift.

Keep-out zones, prohibited areas in Si for processing MOSFETS and other sensitive FEOL devices on IC layouts due to the underfill-microbump stress impact, are proposed. Circular keep-out zones, primarily above the microbump position are proposed for n-type devices. Rectangular keep-out zones around the microbump position, but not directly above the microbump position, are proposed for p-type devices.

Underfill-microbump stress mitigation guidelines are proposed and ranked according to benefiting impact, from highest to lowest impact on stress reduction in Si:

- increasing die thickness
- choosing a low-stress underfill
- decreasing the pitch between microbumps
- decreasing microbump height
- grouping microbumps in larger arrays

The underfill thermal shrikage defined by its CTE below and above its glass transition temperature has the highest impact on underfill thermo-mechanical behavior. Choosing an underfill with lower thermal shrinkage decreases stress levels in Si significantly. Additionaly, choosing an underfill with higher glass transition temperature and lower cure shrinkage could reduce stress levels in Si further. A process involving creation of etched rings in Si around the microbump is proposed in attempt to provide a potential alternative to Si stress reduction. When grouping microbumps in an array and when microbump are at sufficiently close pitch for their stress fields to interact, an edge effect can be expected. The microbump array edge effect involves stress peaks in Si observed above microbumps on the edge of the microbump array. Otherwise, increasing the size of the microbump array decreases stress in Si above the microbumps within the microbump array.

MOSFETs were used as global and local stress sensors in 3D stacks. Stress was analyzed on in total three 3D IC stack generations. Si stress was assessed in the following locations on the 3D stack:

- on the opposite side of the microbump
- below a microbump
- between microbumps
- beneath a Cu pad
- combinations of a microbump and a Cu pad

On the opposite side of the microbump, tensile in-plane stress components in Si are dominant. Compressive out-of-plane stress becomes dominant in Si below the microbump. Tensile in-plane stresses and tensile out-of-plane stresses of the same magnitude are observed in Si between microbumps. Between microbumps, but on the opposite Si side of the microbumps out-of-plane stress is negligible while in-plane stress is of the same magnitude as in-plane stress in Si on the microbump side.

The dominant tensile in-plane stress components on the opposite side of the microbump are of the same sign an directly above the microbump projected center, of equal value. P-type transistors in a [110]/[-110]/[001] reference frame will exhibit lower current shifts when placed directly above the projected microbump center as

in that reference frame they have in-plane piezocoefficients of similar value and opposite sign. N-type transistors are preferable for monitoring the underfill-microbump stress mechanism in a [110]/[-110]/[001] reference frame on the opposite side of the microbump as their in-plane piezocoefficients are of the same sign. N-type transistors will therefore not cancel out the impact of in-plane stress components in Si on the opposite side of the microbump.

In Si below the microbump, n-type transistors could be more desirable as they exhibit a high sensitivity to out-of-plane stress. However, a combination of n-type and p-type sensors below the microbump enables individual in-plane and out-of-plane stress component extraction.

When stress in Si on the opposite side of the microbump is observed through the stack generations, a stress decrease is successfully observed, from an initial value of 827 MPa down to 150 MPa on the FUJI stacks and 125 MPa on the PTCQ stacks.

Out-of-plane stress extracted below the microbump equals to close to 400 MPa of compressive stress. In-plane and out-of-plane stress between the microbumps is below 100 MPa with out-of-plane stress lower than in-plane stress. The Cu pad induces non-negligible negative current shifts on n-type transistors below. Plastic deformation of Cu, most likely during thermo-compression bonding of the stack, could be the reason for remaining stress in Si below the Cu pad.

Current shift decrease due to stress relaxation at higher temperatures was observed. Linear extrapolation of current shifts refered to an equivalent stress free temperature of the 3D IC stack just above 200 ºC. Finite element simulations show good agreement with observed current shifts at room temperature if the stack thermo-mechanical bonding temperature of 250 ºC is taken into account. According to electrical measurements on stress sensors and finite element simulations, stress build up above the glass transition temperature should not be neglected. More complex finite element models need to be built, based on more detailed studies of underfill behaviour above its T_g in order to extract the real equivalent zero stress temperature of a 3D stack.

9.1.4 Analysis of stress in 3D stacked IC packages

Stress is analyzed in 2D and 3D SIC packages. MOSFETs were used as global and local stress sensors. In total 3 package types were assembled:

- 2D BGA package with internal wirebonds
- 2D BGA package with internal Cu pillars
- 3D BGA package

In-plane stresses in Si were extracted in the 2D BGA package with internal wirebonds. The FEOL of the Si die was in this case facing the mold compound. Compressive stress around 120 MPa was observed. High stress is observed in random locations on the die possibly linked to mold compound filler particles pressing on the die during mold compound cooling. Convex warpage of the package surface was observed with optical and mechanical profilometers. X-ray revealed concave warpage of the inner Si die. The convex shape of the package is linked to the mold compound cure shrinkage during mold compound hardening. The concave shape of the die is linked to the

CTE mismatch between the mold compound and the substrate. After the convex shape of the mold compound is created due to mold compound cure shrinkage at the mold compound processing temperature of 180 ºC, during cooling of the package to room temperature, due to higher thermal shrinkage of the mold compound, concave warpage is emposed on the stack. Convex warpage remains on the outer side of the package above the die, while concave warpage impact is observed on the package edges further from the die where only mold compound and substrate interact.

In-plane and out-of-plane stresses in Si were extracted in the 2D BGA package with Cu pillars. The FEOL of the Si die was in this case facing the Cu pillars and substrate. Compressive in-plane stress and out-of-plane stress were observed after die bonding to the substrate, without present mold compound. After additional processing of the mold compound, compressive, but lower in-plane stress and slightly tensile out-of-plane stress is observed.

The die to substrate structure exhibited convex die warpage. It is assumed that during cooling of the die to substrate structure, the substrate with higher CTE than Si shrinks and causes a convex warpage of the structure. Since the die FEOL is on the substrate side, convex warpage on the back Si die surface points to tensile stress on the back Si die surface and compressive stress on the inner Si die surface, where the die FEOL is. When the mold compound is applied, the higher thermal shrinkage of the mold compound compared to the substrate, as in the case of the 2D BGA package with internal wirebonds, attempts to create concave warpage, opposite of the effect of substrate assembly. However, in the case of the 2D BGA with Cu pillars, the die is attached with the substrate by thermo-compression bonding at 250 ºC, higher than the mold compound processing temperature. This means that at mold compound processing temperature of 180 ºC, the die to substrate structure is probably already bent in a convex shape and the mold compound does not succeed in full correction or change of the warpage. This is why the in-plane stresses are still compressive after adding the mold compound but lower than in the die to substrate structure case. The substrate can therefore have a stronger impact on stress in Si than the mold compound. Furthermore, the mold compound could be used to correct the package warpage and to an extent lower Si stress.

In-plane and out-of-plane stresses in Si were extracted from 3D IC packages from both Si tiers, from local and global sensors. Flip chip solder balls were used instead of Cu pillars. Regarding global sensors, placed between the microbumps, on tier 2, the bottom die in the package with its FEOL facing the substrate, compressive in-plane stress around 100 MPa and tensile out-of-plane stress below 100 MPa was extracted. Optical and mechanical profilometry confirm a convex warpage of the 3D IC stack when attached to the substrate. This situation is comparable to the 2D BGA package with Cu pillars, where the FEOL is also facing downwards towards the substrate.

Different stress trends are observed on tier 1, the die further from the substrate, on the mold compound side. This die FEOL is facing the underfill-microbump region and the die it is stacked on, while the back side of the die is turned to the mold compound. There, in the FEOL of tier 1, tensile in-plane stress and higher tensile out-of-plane stress than on tier 2 occurs. It is known from 2D BGA packages, that the mold compound tries to bend the package in a manner opposite of the substrate due to the higher thermal shrinkage of the mold compound compared to the substrate. It can be argued

that the mold compound has a higher and more direct impact on the die further away from the substrate and therefore manages to create concave warpage of the top Si die resulting in such stress fields. Although the substrate undoubtably plays a role in the overall warpage, the mold compound does not interact only with the substrate but with the bottom Si die of the 3D IC stack as well. In one approximation, the bottom Si die can be viewed as the substrate. In terms of warpage of tier 1, in that case, two interactions are occuring. The mold compound to substrate interaction and the mold compound to the bottom die interaction.

Local sensors were also analyzed in the 3D IC package. Current shift shapes observed from the underfill-microbump mechanism are preserved and only globally shifted after packaging. The stress shapes and local current shifts on tier 1 are preserved regardless of the mold compound impact observed on the global sensors as the out-of-plane stress caused by the underfill-microbump mechanism plays a dominant role.

9.2 Recommendations for future work

In chapter 5, stress calibration of FEOL devices was discussed. 4-point bending provides a controlled and straightforward way for application of in-plane stress to FEOL devices. However, a more straightforward method for application of out-of-plane stress can be developed. The usage of a nano-indenter spherical tip for out-of-plane stress calibration makes it difficult to control the distribution of stress to the FEOL due to the tip curvature an unknown contact surface. Out-of-plane stress calibration with help of a flat Si cube provides a more homogenous stress, however the current procedure is not fit for calibration on large number of samples. A more user-friendly approach based on the Si cube or other can be though of.

Furthermore, development of a shear tool for calibration of the FEOL devices to shear stress would be of importance. If developed, FEOL devices could then be calibrated to all stress components and their sensitivity to stress fully characterized. The local effect of BEOL structures during out-of-plane calibration of the FEOL can further be investigated.

In chapter 6, BEOL capacitive plates were analyzed as out-of-plane stress sensors. Using the CIEF CBCM method, stable capacitance was extracted and the device showed sensitivity to out-of-plane stress in a controled environment. However, when used in 3D stacks and 2D or 3D packages, it displayed high instability of capacitance and often inconsistent trends. More study on the operation and applicability of this type of sensors is needed. Furthermore, capacitive plates could be considered also as shear stress sensors.

In chapter 7, stress in 3D SICs was analyzed. The underfill is the generator of stress in 3D stacks and while its impact at room temperature is obvious, its behavior at elevated temperatures, in particular above its T_g, is not well known. Better understanding of Si stress build up in 3D stacks above underfill T_g would possibly lead to the extraction of the equivalent zero stress temperature of the 3D SIC. This value together with a more complex underfill model would be invaluable for stress predictions using finite element modeling. Furthermore, other 3D stack dimensions or material proper-

ties could be proposed for Si stress reduction. New structures, such as the Si etch ring could be further proposed. Addition of materials with a low Young's modulus between the underfill and Si BEOL or Si backside passivation, depending on die orientation, could perhaps act as a stress buffer layer and contribute to stress reduction in Si.

In order to ease 3D SIC design and IC layout design, a compact mechanical model relating 3D stack dimensional and material characteristics would be of great use to the designer. Such a compact mechanical model would incorporate the main learnings from previous finite element simulations and experiments while providing a simpler interface for the designer not directly familiar with thermo-mechanical behavior of 3D SICs and reducing design time.

Also, more fundamental understanding on the creation of tensile stresses in Si in between microbumps is needed.

In chapter 8, stress after packaging 3D SICs was analyzed. While the underfill-microbump interaction stress patterns were still present, the global stress on the two Si dies differed in sign. More fundamental understanding on the change in sign of in-plane stress from one die to the other die and the presence of high out-of-plane stress is needed.

Bibliography

[1] "http://www.radiomuseum.org/forum/usa_project_tinkertoy.html," 2013.

[2] "http://www.chipsetc.com/the-rca-micromodule.html," 2013.

[3] "http://www.computerhistory.org/semiconductor/timeline/1964-hybrid.html," 2013.

[4] Y. Akasaka, "Three-dimensional IC trends," *Proceedings of the IEEE*, vol. 74, no. 12, 1986.

[5] S. Al-Sarawi, D. Abbott, and P. Franzon, "A review of 3-D packaging technology," *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, vol. 21, no. 1, pp. 2–14, 1998.

[6] K. Banerjee, S. Souri, P. Kapur, and K. Saraswat, "3-D ICs: a novel chip design for improving deep-submicrometer interconnect performance and systems-on-chip integration," *Proceedings of the IEEE*, vol. 89, pp. 602–633, May 2001.

[7] S. Wong, A. El-Gamal, and P. Griffin, "Monolithic 3D integrated circuits," in *International Symposium on VLSI Technology, Systems and Applications*, no. c, pp. 1–4, 2007.

[8] P. Batude, M. Vinet, A. Pouydebasque, C. Le Royer, B. Previtali, C. Tabone, J.-M. Hartmann, L. Sanchez, L. Baud, V. Carron, A. Toffoli, F. Allain, V. Mazzocchi, D. Lafond, S. Deleonibus, and O. Faynot, "3D monolithic integration," in *International Symposium of Circuits and Systems*, pp. 2233–2236, Ieee, May 2011.

[9] H. Kurino, K. Lee, T. Nakamura, K. Sakuma, K. Park, N. Miyakawa, H. Shimazutsu, K. Kim, K. Inamura, and M. Koyanagi, "Intelligent image sensor chip with three dimensional structure," in *International Electron Devices Meeting*, vol. 6, pp. 879–882, Ieee, 1999.

[10] E. Beyne, "Technologies for very high bandwidth electrical interconnects between next generation VLSI circuits," in *International Electron Devices Meeting*, pp. 23.3.1–23.3.4, Ieee, 2001.

[11] T. Loher and D. Schutze, "Module miniaturization by ultra thin package stacking," in *Electronic System-Integration Technology Conference*, pp. 1–5, 2010.

[12] C. Bartlett, J. Segelken, and N. Teneketges, "Multichip Packaging Design for VLSI-Based Systems," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, vol. 10, pp. 647–653, Dec. 1987.

[13] A. Fontanelli, "System-in-package technology: opportunities and challenges," in *International Symposium on Quality Electronic Design*, pp. 589–593, 2008.

[14] K. Gann, "Neo-Stacking of Packaged Flash Memory," in *Irvine Sensors corporation*, 2001.

[15] J. Miettinen, M. Mantysalo, K. Kaija, and E. Ristolainen, "System design issues for 3D system-in-package (SiP)," in *Electronic Components and Technology Conference*, pp. 610–615, Ieee, 2004.

[16] "http://www.spacephotonics.com/," 2013.

[17] "http://www.monolithic3d.com/," 2013.

[18] S. S. Wong and A. E. Gamal, "The prospect of 3D-IC," in *Custom Integrated Circuits Conference*, no. Cicc, pp. 445–448, Ieee, Sept. 2009.

[19] B. Black, D. Nelson, C. Webb, and N. Samra, "3D processing technology and its impact on iA32 microprocessors," in *International Conference on Computer Design*, pp. 316–318, Ieee, 2004.

[20] M. Gonzalez, R. Labie, B. Vandevelde, and E. Beyne, "Influence of dielectric materials and via geometry on the thermomechanical behavior of silicon through interconnects," in *Proceedings of Pan Pacific Symposium*, pp. 1–6, 2005.

[21] C. Okoro, K. Vanstreels, R. Labie, O. Lühn, B. Vandevelde, B. Verlinden, and D. Vandepitte, "Influence of annealing conditions on the mechanical and microstructural behavior of electroplated Cu-TSV," *Journal of Micromechanics and Microengineering*, vol. 20, pp. 1–6, Apr. 2010.

[22] G. V. D. Plas and P. Limaye, "Design issues and considerations for low-cost 3-D TSV IC technology," in *International Solid-State Circuits Conference*, vol. 1, pp. 1–3, 2011.

[23] I. De Wolf, K. Croes, O. Varela Pedreira, R. Labie, A. Redolfi, M. Van De Peer, K. Vanstreels, C. Okoro, B. Vandevelde, and E. Beyne, "Cu pumping in TSVs: Effect of pre-CMP thermal budget," *Journal of Microelectronics Reliability*, vol. 51, pp. 1856–1859, Sept. 2011.

[24] B. Swinnen and W. Ruythooren, "3D integration by Cu-Cu thermo-compression bonding of extremely thinned bulk-Si die containing 10 μm pitch through-Si vias," *International Electron Devices Meeting*, pp. 3–6, 2006.

[25] B. Vandevelde and C. Okoro, "Thermo-mechanics of 3D-wafer level and 3D stacked IC packaging technologies," *International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Micro-Systems*, no. April, pp. 21–24, 2008.

[26] J. Van Olmen, A. Mercha, G. Katti, C. Huyghebaert, J. Van Aelst, E. Seppala, Z. Chao, S. Armini, J. Vaes, R. C. Teixeira, M. Van Cauwenberghe, P. Verdonck, K. Verhemeldonck, A. Jourdain, W. Ruythooren, M. de Potter de ten Broeck, A. Opdebeeck, T. Chiarella, B. Parvais, I. Debusschere, T. Hoffmann, B. De Wachter, W. Dehaene, M. Stucchi, M. Rakowski, P. Soussan, R. Cartuyvels, E. Beyne, S. Biesemans, and B. Swinnen, "3D stacked IC demonstration using a through Silicon Via First approach," in *International Electron Devices Meeting*, pp. 1–4, Ieee, Dec. 2008.

[27] C. Okoro, R. Agarwal, P. Limaye, B. Vandevelde, D. Vandepitte, and E. Beyne, "Insertion bonding: A novel Cu-Cu bonding approach for 3D integration," in *Electronic Components and Technology Conference*, pp. 1370–1375, Ieee, 2010.

[28] K. H. Lu, S.-K. Ryu, Q. Zhao, X. Zhang, J. Im, R. Huang, and P. S. Ho, "Thermal stress induced delamination of through silicon vias in 3-D interconnects," *2010 Proceedings 60th Electronic Components and Technology Conference (ECTC)*, pp. 40–45, 2010.

[29] T. Jiang, C. Wu, L. Spinella, J. Im, N. Tamura, M. Kunz, H.-Y. Son, B. Gyu Kim, R. Huang, and P. S. Ho, "Plasticity mechanism for copper extrusion in through-silicon vias for three-dimensional interconnects," *Applied Physics Letters*, vol. 103, no. 21, p. 211906, 2013.

[30] a. Heryanto, W. Putra, a. Trigg, S. Gao, W. Kwon, F. Che, X. Ang, J. Wei, R. I Made, C. Gan, and K. Pey, "Effect of Copper TSV Annealing on Via Protrusion for TSV Wafer Fabrication," *Journal of Electronic Materials*, vol. 41, pp. 2533–2542, May 2012.

[31] F. X. Che, W. N. Putra, A. Heryanto, A. Trigg, S. Member, X. Zhang, S. Member, and C. L. Gan, "Study on Cu Protrusion of Through-Silicon Via," vol. 3, no. 5, pp. 732–739, 2013.

[32] C. McDonough and B. Backes, "Thermal and spatial profiling of TSV-induced stress in 3DICs," in *International Reliability Physics Symposium*, no. 11, pp. 527–532, 2011.

[33] K.-H. Lu, S.-K. Ryu, J. Im, R. Huang, and P. S. Ho, "Thermomechanical reliability of through-silicon vias in 3D interconnects," in *International Reliability Physics Symposium*, pp. 3D.1.1–3D.1.7, Ieee, Apr. 2011.

[34] W. Guo, G. Van der Plas, A. Ivankovic, V. Cherman, G. Eneman, B. De Wachter, M. Togo, A. Redolfi, S. Kubicek, Y. Civale, T. Chiarella, B. Vandevelde, K. Croes, I. De Wolf, I. Debusschere, A. Mercha, A. Thean, G. Beyer, B. Swinnen, and E. Beyne, "Impact of through silicon via induced mechanical stress on fully depleted Bulk FinFET technology," in *International Electron Devices Meeting*, pp. 18.4.1–18.4.4, Ieee, Dec. 2012.

[35] W. Guo, V. Moroz, G. Van der Plas, A. Ivankovic, M. Choi, A. Redolfi, L. Smith, G. Eneman, S. Van Huylenbroeck, P. D. Su, B. De Wachter, I. Debusschere,

K. Croes, I. De Wolf, A. Mercha, G. Beyer, B. Swinnen, and E. Beyne, "Copper through silicon via induced keep out zone for 10nm node bulk FinFET CMOS technology," in *International Electron Devices Meeting*, pp. 12.8.1–12.8.4, Ieee, Dec. 2013.

[36] B. Swinnen and W. Ruythooren, "3D integration by Cu-Cu thermo-compression bonding of extremely thinned bulk-Si die containing 10 μm pitch through-Si vias," in *International Electron Devices Meeting*, pp. 3–6, 2006.

[37] C. Okoro and M. Gonzalez, "Analysis of the induced stresses in silicon during thermcompression Cu-Cu bonding of Cu-through-vias in 3D-SIC architecture," in *Electronic Components and Technology Conference*, pp. 249–255, 2007.

[38] C. Okoro and M. Gonzalez, "Prediction of the influence of induced stresses in silicon on CMOS performance in a Cu-through-via interconnect technology," in *International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Micro-Systems*, pp. 1–7, 2007.

[39] B. Vandevelde and C. Okoro, "Thermo-mechanics of 3D-wafer level and 3D stacked IC packaging technologies," in *International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Micro-Systems*, no. April, pp. 21–24, 2008.

[40] C. Huyghebaert, J. Van Olmen, O. Chukwudi, J. Coenen, A. Jourdain, M. Van Cauwenberghe, R. Agarwahl, A. Phommahaxay, M. Stucchi, and P. Soussan, "Enabling 10um pitch hybrid Cu-Cu IC stacking with Through Silicon Vias," in *Electronic Components and Technology Conference*, pp. 1083–1087, Ieee, 2010.

[41] K. Sakuma, P. Andry, B. Dang, J. Maria, C. Tsang, C. Patel, S. Wright, B. Webb, E. Sprogis, S. Kang, R. Polastre, R. Horton, and J. Knickerbocker, "3D Chip Stacking Technology with Low-Volume Lead-Free Interconnections," in *Electronic Components and Technology Conference*, vol. L, pp. 627–632, Ieee, 2007.

[42] R. Agarwal and W. Zhang, "High density Cu-Sn TLP bonding for 3D integration," in *Electronic Components and Technology Conference*, vol. 32, pp. 345–349, 2009.

[43] R. Agarwal, W. Zhang, and P. Limaye, "Cu/Sn microbumps interconnect for 3D TSV chip stacking," in *Electronic Components and Technology Conference*, vol. 1555, pp. 858–863, 2010.

[44] C. Zhai, U. Ozkan, and A. Dubey, "Investigation of Cu/low-k film delamination in flip chip packages," in *Electronic Components and Technology Conference*, pp. 709–717, 2006.

[45] C. J. Uchibori, X. Zhang, P. S. Ho, T. Nakamura, F. Labs, and E. A. Ave, "Impact of Chip-Package Interaction on Reliability of Cu / Ultra low-k Interconnects for 65nm Technology and beyond," in *International Conference on Solid-State and Integrated Circuit Technology*, no. 408, pp. 5–8, 2006.

[46] C. J. Uchibori, P. S. Ho, and T. Nakamura, "Chip Package Interaction and mechanical reliability impact on Cu/ultra low-k interconnects in Flip Chip package," in *International Conference on Solid-State and Integrated-Circuit Technology*, pp. 1219–1222, Ieee, Oct. 2008.

[47] V. Fiori, S. Gallois-Garreignot, C. Tavernier, H. Jaouen, and A. Juge, "Chip-package interactions: Some combined package effects on copper/low-k interconnect delaminations," in *Electronics System-Integration Technology Conference*, pp. 713–718, Ieee, Sept. 2008.

[48] H.-P. Wei, H.-Y. Tsai, Y.-W. Liu, H.-W. Chen, S.-P. Jeng, and D. C. Yu, "Chip-packaging interaction in Cu/very low-k interconnect," in *International Interconnect Technology Conference*, pp. 143–145, Ieee, June 2009.

[49] L. Fu, M. Su, A. Anand, E. Goh, and F. Kuechenmeister, "Chip package interaction (CPI) reliability of low-k/ULK interconnect with lead free technology," in *Electronic Components and Technology Conference*, pp. 1613–1617, 2010.

[50] J. Auersperg, D. Vogel, M. Lehr, M. Grillberger, and B. Michel, "Crack and damage in low-k BEoL stacks under assembly and CPI aspects," in *International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems*, pp. 1–6, Ieee, Apr. 2010.

[51] M. Chin and A. Marathe, "Reliability Evaluation Methodology for Chip-Package Interaction," in *International Reliability Physics Symposium*, pp. 453–457, 2010.

[52] H. Geisler and M. Lehr, "CPI assessment using a novel characterization technique based on bump-assisted scratch-indentation testing," in *Interconnect Technology Conference*, pp. 1–3, 2011.

[53] D. Chen, W. Chiou, M. Chen, T. Wang, K. Ching, H. Tu, W. Wu, C. Yu, K. Yang, H. Chang, M. Tseng, C. Hsiao, Y. Lu, H. Hu, Y. Lin, C. Hsu, W. S. Shue, and C. Yu, "Enabling 3D-IC foundry technologies for 28 nm node and beyond: through-silicon-via integration with high throughput die-to-wafer stacking," in *International Electron Devices Meeting*, pp. 1–4, Ieee, Dec. 2009.

[54] A. Yu, J. H. Lau, S. W. Ho, A. Kumar, W. Y. Hnin, D.-Q. Yu, M. C. Jong, V. Kripesh, and D. Pinjala, "Study of 15um pitch solder microbumps for 3D IC integration," in *Electronic Components and Technology Conference*, pp. 6–10, Ieee, May 2009.

[55] S. Kohara, K. Sakuma, Y. Takahashi, T. Aoki, K. Matsumoto, P. S. Andry, C. K. Tsang, E. J. Sprogis, J. U. Knickerbocker, and Y. Orii, "Thermal Stress Analysis of 3D Die Stacks with Low-Volume Interconnections," in *CPMT Symposium Japan*, vol. I, pp. 1–4, 2010.

[56] C.-J. Zhan, C.-C. Chuang, J.-Y. Juang, S.-T. Lu, and T.-C. Chang, "Assembly and reliability characterization of 3D chip stacking with 30um pitch lead-free

solder micro bump interconnection," in *Electronic Components and Technology Conference*, pp. 1043–1049, Ieee, 2010.

[57] C.-k. Lee, C.-j. Zhan, J. H. Lau, Y.-j. Huang, H.-c. Fu, J.-h. Huang, Z.-c. Hsiao, S.-y. Huang, C.-w. Fan, Y.-m. Lin, K.-s. Kao, C.-t. Ko, T.-h. Chen, R. Lo, and M. J. Kao, "Wafer Bumping , Assembly , and Reliability Assessment of ubumps with 5um Pads on 10um Pitch for 3D IC Integration," in *IEEE Transactions on Components, Packaging and Manufacturing Technology*, pp. 636–640, 2012.

[58] D. H. Woo, D. L. Lewis, and H.-h. S. Lee, "An Optimized 3D-Stacked Memory Architecture by Exploiting Excessive, High-Density TSV Bandwidth," in *International Symposium on High Performance Computer Architecture*, 2009.

[59] X. Wu, W. Zhao, M. Nakamoto, C. Nimmagadda, D. Lisk, S. Gu, R. Radojcic, M. Nowak, and Y. Xie, "Electrical Characterization for Interier Connections and Timing Analysis for 3-D ICs," *IEEE Transactions on Very Large Scale Integration*, vol. 20, no. 1, pp. 187–190, 2012.

[60] Z. Zhang, B. Noia, K. Chakrabarty, and P. Franzon, "Face-to-face bus design with built-in self-test in 3D ICs," in *International 3D Systems Integration Conference*, pp. 1–7, Ieee, Oct. 2013.

[61] E. J. Marinissen, "Testing TSV-based three-dimensional stacked ICs," in *Design, Automation & Test in Europe Conference & Exhibition*, pp. 1689–1694, Ieee, Mar. 2010.

[62] E. J. Marinissen, S. Dc, and S. Wats, "A Structured and Scalable Test Access Architecture for TSV-Based 3D Stacked ICs," in *IEEE VLSI Test Symposium*, pp. 269–274, 2010.

[63] D. Agonafer, A. Kaisare, M. M. Hossain, Y. Lee, B. P. Dewan-Sandur, T. Dishongh, and S. Pekin, "Thermo-Mechanical Challenges in Stacked Packaging," *Heat Transfer Engineering*, vol. 29, pp. 134–148, Feb. 2008.

[64] S. Melamed and T. Thorolfsson, "Investigation of tier-swapping to improve the thermal profile of memory-on-logic 3DICs," in *International Workshop on Thermal Investigations of ICs and Systems*, pp. 1–6, 2010.

[65] A. Jain, R. Jones, R. Chatterjee, and S. Pozder, "Analytical and Numerical Modeling of the Thermal Performance of Three-Dimensional Integrated Circuits," *IEEE Transactions on Components and Packaging Technologies*, vol. 33, pp. 56–63, Mar. 2010.

[66] V. Venkatadri, B. Sammakia, K. Srihari, and D. Santos, "A Review of Recent Advances in Thermal Management in Three Dimensional Chip Stacks in Electronic Systems," *Journal of Electronic Packaging*, vol. 133, no. 4, pp. 041011–041015, 2011.

[67] H. Oprins, V. Cherman, K. Rebibis, K. Vermeersch, C. Gerets, B. Vandevelde, A. L. Manna, G. Beyer, and E. Beyne, "Transient Analysis Based Thermal Characterization of Die-die Interfaces in 3D-ICs," in *Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, pp. 1395–1404, 2012.

[68] H. Oprins, V. O. Cherman, B. Vandevelde, G. Van der Plas, P. Marchal, and E. Beyne, "Numerical and experimental characterization of the thermal behavior of a packaged DRAM-on-logic stack," in *Electronic Components and Technology Conference*, pp. 1081–1088, Ieee, May 2012.

[69] J. Knickerbocker, P. Andry, B. Dang, R. Horton, C. S. Patel, R. Polastre, K. Sakuma, E. Sprogis, C. Tsang, B. Webb, and S. Wright, "3D silicon integration," in *Electronic Components and Technology Conference*, pp. 538–543, Ieee, May 2008.

[70] G. Van der Plas, S. Thijs, D. Linten, G. Katti, P. Limaye, A. Mercha, M. Stucchi, H. Oprins, B. Vandevelde, N. Minas, M. Cupac, M. Dehan, M. Nelis, R. Agarwal, W. Dehaene, Y. Travaly, E. Beyne, and P. Marchal, "Verifying electrical/thermal/thermo-mechanical behavior of a 3D stack - Challenges and solutions," in *IEEE Custom Integrated Circuits Conference 2010*, pp. 1–4, Ieee, Sept. 2010.

[71] S. W. Yoon, D. W. Yang, J. H. Koo, M. Padmanathan, and F. Carson, "3D TSV processes and its assembly/packaging technology," in *International Conference on 3D System Integration*, pp. 1–5, Ieee, Sept. 2009.

[72] M.-c. Hsieh, C.-k. Yu, and S.-t. Wu, "Thermo-Mechanical Simulative Study for 3D Vertical Stacked IC Packages with Spacer Structures," in *IEEE Semiconductor Thermal Measurement and Management Symposium*, no. 195, pp. 47–54, 2010.

[73] J. Lau, X. Zheng, and C. Selvanayagam, "Failure analyses of 3D Sip (system-in-package) and WLP (wafer-level package) by finite element methods," in *International Symposium on Physical and Failure Analysis of Integrated Circuits*, pp. 1–7, 2009.

[74] C. Stapleton, "Nonlinear metal resistors," *Electonics Letters*, vol. 7, no. 2, pp. 2–3, 1971.

[75] E. Klokholm and A. Mayadas, "Size Effects and Piezoresistance in Metal Films," *Journal of Vacuum Science and . . .*, vol. 9, no. 1, pp. 441–443, 1972.

[76] C. Hu, Y. Gao, and Z. Sheng, "The piezoresistance coefficients of copper and copper-nickel alloys," *Journal of materials science*, vol. 5, no. 5, pp. 381–386, 2000.

[77] J. Terry and S. Smith, "Test chip for the development and evaluation of sensors for measuring stress in metal interconnects," *International Conference on Microelectronic Test Structures*, vol. 18, no. 2, pp. 255–261, 2005.

[78] O. Hansen, K. Reck, and E. V. Thomsen, "Giant geometrically amplified piezoresistance in metal-semiconductor hybrid resistors," *Journal of Applied Physics*, vol. 104, no. 11, p. 114510, 2008.

[79] V. Chandrasekharan, J. Sells, J. Meloy, D. P. Arnold, M. Sheplak, and I. M. Group, "A metal-on-silicon differential capacitive shear stress sensor," in *International Conference on Solid-State Sensors, Actuators and Microsystems*, no. Mask 2, pp. 1537–1540, 2009.

[80] T. H. et al. H. Namba, "Matching Characteristics of Metal Resistors," in *International Conference on Microelectronic Test Structures*, pp. 189–193, 2011.

[81] F. Fruett, *The Piezojunction Effect in Silicon, its Consequences and Applications for Integrated Circuits and Sensors*. 2001.

[82] R. Beaty and J. Suhling, "Calibration considerations for piezoresistive-based stress sensors," in *Electronic Components and Technology Conference*, pp. 797–806, 1990.

[83] J. Suhling and R. Jaeger, "Silicon piezoresistive stress sensors and their application in electronic packaging," *IEEE Sensors Journal*, vol. 1, pp. 14–30, June 2001.

[84] a. Mian, J. Suhling, and R. Jaeger, "The van der Pauw stress sensor," *IEEE Sensors Journal*, vol. 6, pp. 340–356, Apr. 2006.

[85] T. Chen and Y. Huang, "Evaluation of MOS devices as mechanical stress sensors," *IEEE Transactions on Components and Packaging Technology*, vol. 25, no. 3, pp. 511–517, 2002.

[86] A. T. Bradley, R. C. Jaeger, J. C. Suhling, K. J. O. Connor, and S. Member, "Piezoresistive Characteristics of Short-Channel MOSFETs on (100) Silicon," *IEEE Transactions on Electron Devices*, vol. 48, no. 9, pp. 2009–2015, 2009.

[87] R. Jaeger, J. Suhling, R. Ramani, and a.T. Bradley, "CMOS stress sensors on [100] silicon," *IEEE Journal of Solid-State Circuits*, vol. 35, pp. 85–95, Jan. 2000.

[88] Y. Chen, R. C. Jaeger, and J. C. Suhling, "Multiplexed CMOS Sensor Arrays for Die Stress Mapping," in *European Solid-State Circuits Conference*, no. 100, pp. 424–427, 2006.

[89] C. Djelassi, "Comparison of two calibration methods for a package stress measurement testchip," in *International Physic Reliability Symposium*, no. 4, pp. 446–452, 2010.

[90] M. Doelle, C. Peters, P. Ruther, and O. Paul, "Piezo-FET stress-sensor arrays for wire-bonding characterization," *Journal of Microelectromechanical Systems*, vol. 15, no. 1, pp. 120–130, 2006.

[91] M. Doelle, J. Held, P. Ruther, and O. Paul, "Simultaneous and Independent Measurement of Stress and Temperature Using a Single Field-Effect Transistor Structure," *Journal of Microelectromechanical Systems*, vol. 16, pp. 1232–1242, Oct. 2007.

[92] P. Gieschke, B. Sbierski, and O. Paul, "CMOS-based piezo-FET stress sensors in Wheatstone bridge configuration," in *IEEE Sensors Proceedings*, pp. 93–96, Ieee, Oct. 2011.

[93] C. Smith, "Piezoresistance effect in Ge and Si," *Physical review*, vol. 94, no. 1, pp. 42–49, 1954.

[94] R. B. D. Colman and J. Mize, "Mobility Anisotropy and Piezoresistance in Silicon p-Type Inversion Layers," vol. 39, no. 4, 1968.

[95] G. Dorda, "Piezoresistance in Quantized Conduction Bands in Silicon Inversion Layers," *Journal of Applied Physics*, vol. 42, no. 5, pp. 2053–2060, 1971.

[96] Y. Kanda, "A graphical representation of the piezoresistance coefficients in silicon," *IEEE Transactions on Electron Devices*, vol. 29, pp. 64–70, Jan. 1982.

[97] Y. Kanda, "Graphical Representation of the Piezoresistance Coefficients in Silicon - Shear Coefficients in Plane," *Japanese Journal of Applied Physics*, vol. 26, no. 7, pp. 1031–1033, 1987.

[98] Y. Kanda, "Piezoresistance effect of silicon," *Sensors and Actuators A*, vol. 28, pp. 83–91, 1991.

[99] K. Matsuda, K. Suzuki, K. Yamamura, and Y. Kanda, "Nonlinear piezoresistance effects in silicon," *Journal of Applied Physics*, vol. 73, no. 4, p. 1838, 1993.

[100] A. A. Barlian, W.-T. Park, J. R. Mallon, A. J. Rastegar, and B. L. Pruitt, "Review: Semiconductor Piezoresistance for Microsystems.," *Proceedings of the IEEE*, vol. 97, pp. 513–552, Jan. 2009.

[101] B. Lemke, M. E. Schmidt, J. Gutmann, P. Gieschke, P. Alpuim, J. Caspar, and O. Paul, "Nonlinear piezoresistance of silicon," in *IEEE Sensors conference*, pp. 1950–1953, Ieee, Nov. 2010.

[102] J. Gaspar, J. Gutmann, B. Lemke, and O. Paul, "Nonlinear piezoresistance of silicon at large stresses," in *24th International Conference on Micro Electro Mechanical Systems*, pp. 469–472, Ieee, Jan. 2011.

[103] C.-h. Cho, R. C. Jaeger, and J. C. Suhling, "Characterization of the Temperature Dependence of the Piezoresistive Coefficients of Silicon From 150 C to + 125 C," *IEEE Sensors Journal*, vol. 8, no. 8, pp. 1455–1468, 2008.

[104] X.-f. Fan, L. F. Register, S. Member, B. Winstead, M. C. Foisy, W. Chen, X. Zheng, B. Ghosh, and S. K. Banerjee, "Hole Mobility and Thermal Velocity Enhancement for Uniaxial Stress in Si up to 4 GPa," vol. 54, no. 2, pp. 291–296, 2007.

[105] O. Zienkiewicz and R. Taylor, *The Finite Element Method*. McGraw-Hill International (UK), 1989.

[106] N. Ottosen and H. Peterson, *Introduction to the finite element method*. Prentice-Hall, 1992.

[107] "MSC Marc manual," tech. rep., MSC Software, 2013.

[108] "Delaminator manual," tech. rep., Dauskardt Technical Services, 2009.

[109] "MTS Nano Indenter XP," tech. rep., MTS Systems Corporation, 2009.

[110] "DEKTAK XT Stylus manual," tech. rep., Bruker, 2010.

[111] Veeco Metrology Group, "WYKO Surface Profilers - technical reference manual," tech. rep., 1999.

[112] G. Stoney, "The Tension of Metalic Films Deposited by Electrolysis," in *Proceeding of Royal Society A*, pp. 172–175, 1909.

[113] X. Feng, Y. Huang, and A. J. Rosakis, "On the Stoney Formula for a Thin Film/Substrate System With Nonuniform Substrate Thickness," *Journal of Applied Mechanics*, vol. 74, no. 6, pp. 1276–1281, 2007.

[114] W. Friedrich, P. Knipping, and M. Von Laue in *Proc. Bavarian Acad. Sci.*, p. 303, 1912.

[115] W. L. Bragg in *Proceeding of Royal Society A*, p. 248, 1913.

[116] J. Stopford. PhD thesis, Dublin City University, 2011.

[117] J. Stopford, A. Henry, D. Manessis, N. Bennet, K. Horan, D. Allen, J. Wittge, L. Boettcher, A. Cowley, and P. McNally, "Non-destructive X-ray mapping of strain & warpage of die in packaged chips," in *18th European Microelectronics and Packaging Conference*, 2011.

[118] C. S. Wong, *X-ray diffraction techniques for future advanced CMOS metrology challenges*. PhD thesis, Dublin City University, 2013.

[119] P. Gargini, "The global route to future semiconductor technology," *Circuits and Devices Magazine, IEEE*, vol. 18, no. 2, pp. 13–17, 2002.

[120] A. Sedra and K. Smith, *Microelectronic Circuits*. Oxford University Press, 2004.

List of publications

Conference proceedings

1. **A. Ivankovic**, V. Cherman, M. Gonzalez, B. Vandevelde, D. Vandepitte, G. Beyer, E. Beyne, I. De Wolf, "The underfill-microbump interaction mechanism in 3D ICs: impact and mitigation of induced stresses", International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE), pp. 1-8, 2014.

2. V. Cherman, G. Van der Plas, J. De Vos, **A. Ivankovic**, M. Lofrano, M. Gonzalez, K. Vanstreels, T. Wang, R. Daily, W. Guo, G. Beyer, A. La Manna, I. De Wolf, E. Beyne, "3D Stacking Induced Mechanical Stress Effects", Electronics Components and Technology Conference (ECTC), in press, 2014.

3. C.S. Wong, **A. Ivankovic**, A. Cowley, N.S. Bennett, A.N. Danilewsky, M. Gonzalez, V. Cherman, B. Vandevelde, I. De Wolf, P.J. McNally, "Development of B-spline x-ray diffraction imaging techniques for die warpage and stress monitoring inside fully encapsulated packaged chips", Electronics Components and Technology Conference (ECTC), in press, 2014.

4. B. Vandevelde, **A. Ivankovic**, B. Debecker, M. Lofrano, K. Vanstreels, W. Guo, V. Cherman, M. Gonzalez, G. Van der Plas, I. De Wolf, E. Beyne, Z. Tokei, "IC-Package interaction", International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE), pp. 1-4, 2013.

5. W. Guo, V. Moroz, G. Van der Plas, M. Choi, A. Redolfi, L. Smith, G. Eneman, S. Van Huylenbroeck, P.D. Su, **A. Ivankovic**, B. De Wachter, I. Debusschere, K. Croes, I. De Wolf, A. Mercha, G. Beyer, B. Swinnen, E. Beyne, "Copper Through Si Via Induced Keep Out Zone for 10nm Node Bulk FinFET CMOS Technology", International Electron Devices Meeting (IEDM), pp. 12.8.1 - 12.8.4, 2013.

6. B. Vandevelde, **A. Ivankovic**, B. Debecker, M. Lofrano, K. Vanstreels, W. Guo, V. Cherman, M. Gonzalez, G. Van der Plas, I. De Wolf, E. Beyne, Z. Tokei, "Challenges in BEOL Design and Materials coping with CPI Stress issues for Advanced Packaging Solutions", Materials for Advanced Metallization (MAM), pp. 1-4, 2013.

7. **A. Ivankovic**, V. Cherman, G. Van der Plas, B. Vandevelde, G. Beyer, E. Beyne, I. De Wolf, D. Vandepitte, "FET Arrays as CPI Sensors for 3D Stacking and Packaging Characterization", International Reliability Physics Conference (IRPS), pp. 2E.3.1 – 2E.3.9, 2012.

8. M. Gonzalez, B. Vandevelde, **A. Ivankovic**, V. Cherman, B. Debecker, M. Lofrano, I. De Wolf, G. Beyer, B. Swinnen, Z. Tokei, E. Beyne, "Chip Package Interaction (CPI): Thermo Mechanical Challenges in 3D Technologies", Electronics Packaging Technology Conference (EPTC), pp. 547-551, 2012.

9. W.Guo, G. Van der Plas, **A. Ivankovic**, V. Cherman, G. Eneman, B. De Wachter, M. Togo, A. Redolfi, S. Kubicek, Y. Civale, T. Chiarella, B. Vandevelde, K. Croes, I. De Wolf, I. Debusschere, A. Mercha, A. Thean, G. Beyer, B. Swinnen, E. Beyne, "Impact of Through Silicon Via Induced Mechanical Stress of Fully Depleted Bulk FinFET Technology", International Electron Devices Meeting (IEDM), pp. 18.4.1 - 18.4.4, 2012.

10. K. Croes, V. Cherman, Y. Li, L. Zhao, Y. Barbarin, J. De Messemaeker, Y. Civale, D. Velenis, M. Stucchi, T. Kauerauf, A. Redolfi, B. Dimcic, **A. Ivankovic**, G. Van der Plas, I. De Wolf, G. Beyer, B. Swinnen, Z. Tokei, E. Beyne, "Reliability concerns in copper TSVs: Methods and Results", International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), pp. 1-5, 2012.

11. W. Guo, G. Van der Plas, **A. Ivankovic**, G. Eneman, V. Cherman, B. De Wachter, A. Mercha, M. Gonzalez, Y. Civale, A. Redolfi, T. Buisson, A. Jourdan, B. Vandevelde, K.J. Rebibis, I. De Wolf, A. La Manna, G. Beyer, E. Beyne, B. Swinnen, "3D Chip Package Interaction Thermo-Mechanical Challenges: Proximity Effects of Through- Silicon Vias and μ-bumps", International Conference on IC design and Technology (ICICDT), pp. 1-4, 2012.

12. **A. Ivankovic**, B. Vandevelde, K.J. Rebibis, A. La Manna, G. Van der Plas, V. Cherman, E. Beyne, I. De Wolf, D. Vandepitte, "Thermo-mechanical Impact of the Underfill-microbump Interaction in 3D Stacked Integrated Circuits", Electronics Packaging Technology Conference (EPTC), pp. 34-38, 2011.

13. **A. Ivankovic**, G. Van der Plas, V. Moroz, M. Choi, V. Cherman, A. Mercha, P. Marchal, M. Gonzalez, G. Eneman, W. Zhang, T. Buisson, M. Detalle, A. La Manna, D. Verkest, G. Beyer, E. Beyne, B. Vandevelde, I. De Wolf, D. Vandepitte, ''Analysis of Microbump Induced Stress Effects in 3D Stacked IC Technologies'', International 3D System Integration Conference (3DIC), pp. 1 -5, 2011.

14. **A. Ivankovic**, K. Vanstreels, YY Hsu, M. Gonzalez, G. Brizar, D. Vanderstraeten, E. Blansaer, R. Gillon, M. Defloor, K. Vandaele, D. Degryse, B. Vandevelde, "Impact of thermal ageing on cohesive and adhesive strengths of overmold materials: Characterization methods and implementation in FEM", International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE), pp. 1-5, 2011.

Journal papers

1. **A. Ivankovic**, V. Cherman, M. Gonzalez, B. Vandevelde, D. Vandepitte, G. Beyer, E. Beyne, I. De Wolf, "The underfill-microbump interaction mechanism in 3D ICs: impact and mitigation of induced stresses", in submission, 2014.

2. B. Vandevelde, **A. Ivankovic**, B. Debecker, M. Lofrano, K. Vanstreels, W. Guo, V. Cherman, M. Gonzalez, G. Van der Plas, I. De Wolf, E. Beyne, Z. Tokei, "Chip-Package Interaction in 3D stacked IC packages using Finite Element Modeling", Microelectronics Reliability, in press, 2014.

3. L. Kljucar, M. Gonzalez, K. Vanstreels, **A. Ivankovic**, M. Hecker, I. De Wolf, "Effect of 4-point bending test procedure on crack propagation in thin film stacks", Microelectronics Reliability, submitted, 2014.

4. **A. Ivankovic**, K. Vanstreels, D. Vanderstraeten, G. Brizar, R. Gillon, E. Blansaer, B. Vandevelde, "Comparison of experimental methods for the extraction of the elastic modulus of molding compounds used in IC packaging", Microelectronics Reliability, vol. 52, pp. 2677 – 2684, 2012.

Patent contribution

1. W. Zhang, W. Wang, G. Eneman, E. Beyne, **A. Ivankovic**, "Method for Forming Isolation Trenches in Micro-Bump Interconnect Structures and Devices Obtained Thereof", Publication no.: US 2013/0154112, USPTO, 2013.